

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 7月17日
Date of Application:

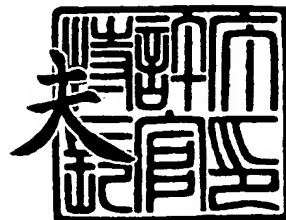
出願番号 特願2003-198707
Application Number:
[ST. 10/C]: [JP2003-198707]

出願人 セイコーエプソン株式会社
Applicant(s):

2003年 8月12日

特許庁長官
Commissioner,
Japan Patent Office

今井 康



出証番号 出証特2003-3065018



【書類名】 特許願

【整理番号】 J0101649

【あて先】 特許庁長官殿

【国際特許分類】 G10L 15/20

【発明者】

 【住所又は居所】 長野県諏訪市大和3丁目3番5号 セイコーエプソン株式会社内

 【氏名】 宮澤 康永

【特許出願人】

 【識別番号】 000002369

 【氏名又は名称】 セイコーエプソン株式会社

【代理人】

 【識別番号】 100095728

 【弁理士】

 【氏名又は名称】 上柳 雅誉

 【連絡先】 0266-52-3528

【選任した代理人】

 【識別番号】 100107076

 【弁理士】

 【氏名又は名称】 藤網 英吉

【選任した代理人】

 【識別番号】 100107261

 【弁理士】

 【氏名又は名称】 須澤 修

【先の出願に基づく優先権主張】

 【出願番号】 特願2002-318627

 【出願日】 平成14年10月31日

【手数料の表示】

【予納台帳番号】 013044

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0109826

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音響モデル作成方法および音声認識装置ならびに音声認識装置を有する乗り物

【特許請求の範囲】

【請求項 1】 雑音を有する空間内で音声認識を行うための音響モデル作成方法であって、

前記雑音を有する空間内で収集可能な各種雑音を収集する雑音収集ステップと

この雑音収集ステップにより収集された雑音を分類して複数種類の雑音データを作成する雑音データ作成ステップと、

この雑音データ作成ステップによって作成された複数種類の雑音データを標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成する雑音重畳音声データ作成ステップと、

この雑音重畳音声データ作成ステップによって作成された複数種類の雑音重畳音声データに対して雑音除去処理を行って複数種類の雑音除去音声データを作成する雑音除去音声データ作成ステップと、

この雑音除去音声データ作成ステップによって作成された複数種類の雑音除去音声データから複数種類の音響モデルを作成する音響モデル作成ステップと、

を有することを特徴とする音響モデル作成方法。

【請求項 2】 前記複数種類の雑音重畳音声データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うことを特徴とする請求項 1 記載の音響モデル作成方法。

【請求項 3】 前記雑音を有するある空間内は、乗り物内であることを特徴とする請求項 1 または 2 記載の音響モデル作成方法。

【請求項 4】 前記乗り物内で収集可能な各種雑音は、天候、前記乗り物の走行状況、前記乗り物の走行位置、前記乗り物に搭載されている機器の動作状態の少なくとも 1 つに起因する複数種類の雑音であることを特徴とする請求項 3 記載の音声認識用の音響モデル作成方法。

【請求項 5】 前記雑音収集ステップは、前記収集すべき複数種類の雑音に

対応するそれぞれの雑音パラメータを記録する雑音パラメータ記録工程を含み、前記雑音データ作成ステップは、前記収集すべき複数種類の雑音とこの収集すべき雑音に対応するそれぞれの雑音パラメータとを用いて分類することによって前記複数種類の雑音データを作成することを特徴とする請求項 1 から 4 のいずれかに記載の音響モデル作成方法。

【請求項 6】 雑音を有する空間内で音声認識を行う音声認識装置であって

認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、

前記雑音を有する空間内で収集可能な各種雑音を収集する雑音収集ステップと、その収集された雑音を分類して複数種類の雑音データを作成する雑音データ作成する雑音データ作成ステップと、その作成された複数種類の雑音データを予め用意された標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成する雑音重畳音声データ作成ステップと、その作成された複数種類の雑音重畳音声データに対して雑音除去処理を行って複数種類の雑音除去音声データを作成する雑音除去音声データ作成ステップと、その作成された複数種類の雑音除去音声データから複数種類の音響モデルを作成する音響モデル作成ステップとを有する音響モデル作成方法により作成された複数種類の音響モデルと、

前記音入力手段に入力された雑音が、前記複数種類の雑音データのどの雑音データに属する雑音であるかを判別する雑音データ判別手段と、

前記音入力手段から雑音の重畳された雑音重畳音声データに対して前記雑音データ判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、

この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記複数種類の音響モデルのうち前記雑音データ判別手段で判別された雑音データに対応した音響モデルを用いて音声認識を行う音声認識手段と、

を有したことを特徴とする音声認識装置。

【請求項 7】 前記音入力手段に入力された雑音に対応する雑音パラメータを取得する雑音パラメータ取得手段を有したことを特徴とする請求項 6 記載の音声認識装置。

【請求項 8】 前記分類によって得られた複数種類の雑音データに対して行

われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うことを特徴とする請求項 6 または 7 記載の音声認識装置。

【請求項 9】 前記雑音を有するある空間内は、乗り物内であることを特徴とする請求項 6 から 8 のいずれかに記載の音声認識装置。

【請求項 10】 前記乗り物内で収集可能な各種雑音は、天候、前記乗り物の走行状況、前記乗り物の走行位置、前記乗り物に搭載されている機器の動作状態の少なくとも 1 つに起因する複数種類の雑音であることを特徴とする請求項 9 記載の音声認識装置。

【請求項 11】 前記音響モデルを作成するための前記雑音収集ステップは、前記収集すべき複数種類の雑音に対応するそれぞれの雑音パラメータを記録する工程を含み、前記雑音データ作成ステップは、前記収集すべき複数種類の雑音とこの収集すべき雑音に対応するそれぞれの雑音パラメータとを用いて分類することによって前記複数種類の雑音データを作成することを特徴とする請求項 6 から 10 のいずれかに記載の音声認識装置。

【請求項 12】 前記複数種類の音響モデルを作成する際の雑音除去処理と前記認識すべき音声に対する音声認識を行う際の雑音除去処理は同じ雑音除去手法を用いることを特徴とする請求項 6 から 11 のいずれかに記載の音声認識装置。

【請求項 13】 雑音を有するある空間内で音声認識を、前記請求項 1 から 5 に記載の音響モデル作成方法によって作成された複数種類の音響モデルを用いて行う音声認識装置であって、

認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、

前記音入力手段から入力された現在の雑音が、予め分類された複数種類の雑音データのどの雑音データに属する雑音であるかを判別する雑音データ判別手段と

、
前記音入力手段から雑音の重畳された雑音重畳音声データに対して前記雑音データ判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、

この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記複数種類の音響モデルのうち前記雑音データ判別手段で判別された雑音種類に対応し

た音響モデルを用いて音声認識を行う音声認識手段と、

を有したことを特徴とする音声認識装置。

【請求項 14】 音声によって機器操作の可能な音声認識装置を有する乗り物であって、

前記音声認識装置として、請求項 6 から 13 のいずれかに記載の音声認識装置を有することを特徴とする音声認識装置を有する乗り物。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、雑音を有する空間内で音声認識を行うための音声認識用の音響モデル作成方法および音声認識装置に関する。また、本発明の音声認識装置を有する乗り物に関する。

【0002】

【従来の技術】

最近、音声認識技術は様々な分野で利用され、各種機器を音声によって操作可能とすることが一般的に行われるようになってきている。このように、音声によってある特定の機器の操作を可能とすることで、両手で何らかの操作を行っているときに、他の機器操作を行う必要のある場合などにおいてきわめて便利なものとなる。

【0003】

たとえば、自動車に搭載されているカーナビゲーション、カーオーディオ、カーエアコンディショナ（以下ではエアコンという）などの様々な機器は、通常は、必要なときに運転者が自らの手を使って操作するのが一般的であるが、最近では、これらの機器を音声で操作可能とする技術が種々提案され実用化されつつある。

これによって、運転中であっても、ステアリングから手を離すことなく、これらの機器のスイッチをオン・オフさせたり、それらの機器の機能設定が可能となるので、安全動作にもつながり、今後、ますます普及して行くものと考えられる。

【0004】

しかし、自動車などに搭載される上述のような機器の操作を音声によって行うには、様々な雑音の存在する環境下で高い認識性能が得られるようにすることが重要であり、この点が従来からの大きな課題ともなっている。

【0005】

このように、自動車の車内など様々な雑音の存在する環境下で音声認識を行う方法として、図15に示すような方法で音響モデルを作成し、その音響モデルを用いて図16に示すように音声認識を行う方法が従来から用いられている。

【0006】

この従来の音声認識方法で用いられる音響モデル作成処理について図15を参照しながら説明する。まず、無響室などの雑音のない環境で収集された標準的な音声データ（たとえば、多数の話者が多種類単語について発話して得られた大量の音声データ） V と、ある特定種類の雑音データ N を雑音重畳データ作成部51に入力して、その標準的な音声データにある特定種類の雑音をある S/N 比で重畳させた雑音重畳音声データ VN を作成する。

【0007】

この雑音重畳音声データ VN に対し、雑音除去処理部52がスペクトラム・サブトラクション（SS）やケプストラム平均化処理（CMN）など、その雑音の種類に最適な雑音除去処理を施し、雑音除去音声データ V' （雑音除去処理を行っても除去されない雑音成分が残る）を作成する。そして、音響モデル学習処理部53がこの雑音除去音声データ V' を用いて音素HMM（Hidden Markov Model）や音節HMMなどの音響モデル M を作成する。

【0008】

一方、この従来の音声認識処理は図16に示すように、マイクロホン61から入力された話者の音声データ（機器操作の音声コマンド）に対し、入力信号処理部62が増幅や A/D 変換（アナログ／デジタル変換）などを行い、そのあと、雑音除去処理部63が入力音声データに対して雑音除去処理（図15の雑音除去処理部52で行ったと同じ手法による雑音除去処理）を行う。

【0009】



そして、その雑音除去された音声データ（雑音除去音声データという）に対し、音声認識処理部 64 が言語モデル 65 と図 8 の音響モデル学習処理部 53 で作成された音響モデル M を用いて音声認識処理する。

【0010】

しかし、上述した従来の音声認識手法では、ある特定の雑音のみに対応して作成された音響モデル M のみを使って音声認識するようにしているので、上述したような刻々と変化する多種多様な雑音には対応できず、状況によって発生する雑音が音声認識性能に大きく影響を与え、高い認識率を得ることは難しいという問題がある。

【0011】

これに対して、特開 2002-132289 号公報に記載の技術のように、様々な雑音の種類に応じた複数種類の音響モデルを作成し、実際の音声認識時には、音声に重畳された雑音に応じて、複数種類の音響モデルの中から最適な音響モデルを選択して音声認識を行うものがある。

【0012】

【特許文献 1】

特開 2002-132289 号公報

【0013】

【発明が解決しようとする課題】

上述した特許文献 1 によれば、いくつかの雑音に対応した音響モデルを有し、そのときの雑音に最適な音響モデルを選択して音声認識することによって、確かに、高精度な音声認識が可能となる。

【0014】

しかし、自動車の車内で音声認識を行う場合、自動車の走行状況に起因する音（速度などに応じたタイヤのパターンノイズや窓の開度に応じた風切り音、回転数や変速ギヤ位置などによるエンジン音）、周囲の環境に起因する音（トンネルなどを通過する際の反響音など）、自動車に搭載されている機器の動作状態に起因する音（カーオーディオの動作音やエアコンの動作音、ワイパーや方向指示器の動作音など）、降雨時における雨音など自動車特有の雑音がマイクロホンから

入り、これらの雑音が音声コマンドに重畳した状態で以降の音声認識処理部に渡されることになる。

【0015】

一般に、自動車の場合、マイクロホンから入力される雑音の種類は上述したように、自動車特有の雑音であり、その種類はある程度限定されるが、走行状況に起因する雑音としてエンジン音だけに注目して考えても、アイドリング時、低速走行時、高速走行時では雑音の大きさや雑音の種類が異なる場合が多い。また。同じ速度で走行していても、変速ギヤ比の関係でエンジンの回転数が高いときと低いときで雑音の大きさや種類が異なるなど、エンジン音だけを考えても、雑音の大きさと種類は様々異なってくる。

【0016】

さらに、このような走行状況に起因するものだけではなく、上述したように、たとえば、窓の開閉度合いによる風切り音、トンネルや橋など周囲の建造物などの反響音、降雨時（降雨量によっても異なる）による雨音、エアコン、ワイパー、カーオーディオ、方向指示器など各種機器の動作音も雑音としてマイクロホンに入力されてくる。

【0017】

このように、自動車で発生する雑音は、雑音種類はある程度は限定されるものの、状況によっては同じ種類の雑音であっても様々異なってくることも大きな特徴であり、このような雑音環境下にあっては、前述の特許文献1の技術では対応できない場合がある。

【0018】

また、これは自動車だけではなく、他の乗り物などでも同じである。さらに、乗り物だけではなく、たとえば、工場や流通センタなどの作業場で音声認識を行う場合も、雑音の種類は自動車とは異なるものの、自動車での音声認識を行う場合と同様のことが言える。

【0019】

そこで本発明は、たとえば、雑音を有する空間内で音声認識する際、その空間内の雑音環境に適した音声認識を可能とするための音響モデルを作成する音響モ



デル作成方法および多種多様な雑音環境下で高い認識性能を得ることのできる音声認識装置、さらに、この音声認識装置を備えることによって、多種多様な雑音環境下においても音声による機器操作などを確実に行うことのできる音声認識装置を有する乗り物を提供することを目的としている。

【0020】

【課題を解決するための手段】

(1) 本発明の音響モデル作成方法は、雑音を有する空間内で音声認識を行うための音響モデル作成方法であって、前記雑音を有する空間内で収集可能な各種雑音を収集する雑音収集ステップと、この雑音収集ステップにより収集された雑音を分類して複数種類の雑音データを作成する雑音データ作成ステップと、この雑音データ作成ステップによって作成された複数種類の雑音データを標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成する雑音重畳音声データ作成ステップと、この雑音重畳音声データ作成ステップによって作成された複数種類の雑音重畳音声データに対して雑音除去処理を行って複数種類の雑音除去音声データを作成する雑音除去音声データ作成ステップと、この雑音除去音声データ作成ステップによって作成された複数種類の雑音除去音声データから複数種類の音響モデルを作成する音響モデル作成ステップとを有することを特徴としている。

【0021】

このように、ある空間内で収集された雑音を分類して複数種類の雑音データを作成し、その複数種類の雑音データを予め用意された標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成し、その複数種類の雑音重畳音声データに対して雑音除去処理を行って、その複数種類の雑音除去音声データから複数種類の音響モデルを作成するようにしているので、その空間内での様々な雑音の種類に対応した最適な音響モデルを作成することができる。

【0022】

(2) 前記(1)の音響モデル作成方法において、複数種類の雑音重畳音声データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うようにしている。

これによって、それぞれの雑音データに対して適正かつ効率的な雑音除去が可能となる。

【0023】

(3) 前記(1)または(2)の音響モデル作成方法において、雑音を有するある空間内は、乗り物内であることが1つの例として考えられる。

これによって、乗り物(たとえば、自動車)特有の様々な雑音の種類に対応した最適な音響モデルを作成することができる。

【0024】

(4) 前記(3)の音響モデル作成方法において、乗り物内で収集可能な各種雑音は、天候、前記乗り物の走行状況、前記乗り物の走行位置、前記乗り物に搭載されている機器の動作状態の少なくとも1つに起因する複数種類の雑音としている。

【0025】

この雑音は、乗り物が自動車である場合、その一例を挙げれば、走行速度などに応じたエンジン音やタイヤのパターンノイズ、降雨時における雨音、エアコンやカーオーディオ機器などの車載機器の動作音などである。そして、これらの音が雑音として収集され、これらの雑音を分類して、それぞれの雑音グループに対応した雑音データを生成し、それぞれの雑音データごとの音響モデルを作成することによって、乗り物、特に自動車特有の様々な雑音に対応可能な音響モデルを作成することができる。

【0026】

(5) 前記(1)から(4)のいずれかの音響モデル作成方法において、前記雑音収集ステップは、前記収集すべき複数種類の雑音に対応するそれぞれの雑音パラメータを記録する雑音パラメータ記録工程を含み、前記雑音データ作成ステップは、前記収集すべき複数種類の雑音とこの収集すべき雑音に対応するそれぞれの雑音パラメータとを用いて分類することによって前記複数種類の雑音データを作成するようにしている。

【0027】

この雑音パラメータは、たとえば、その一例を挙げると、自動車の速度を示す

情報、エンジン回転数を示す情報、エアコンの動作状態を示す情報などである。これらの雑音パラメータを雑音とともに記録することで、たとえば、どのような速度のときはどのような雑音が発生するかの対応付けが行え、適切な分類が可能となり、実際の雑音環境に適した雑音データ得ることができる。

【0028】

(6) 本発明の音声認識装置は、雑音を有する空間内で音声認識を行う音声認識装置であって、認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、前記雑音を有する空間内で収集可能な各種雑音を収集する雑音収集ステップと、その収集された雑音を分類して複数種類の雑音データを作成する雑音データ作成する雑音データ作成ステップと、その作成された複数種類の雑音データを予め用意された標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成する雑音重畳音声データ作成ステップと、その作成された複数種類の雑音重畳音声データに対して雑音除去処理を行って複数種類の雑音除去音声データを作成する雑音除去音声データ作成ステップと、その作成された複数種類の雑音除去音声データから複数種類の音響モデルを作成する音響モデル作成ステップとを有する音響モデル作成方法により作成された複数種類の音響モデルと、前記音入力手段に入力された雑音が、前記複数種類の雑音データのどの雑音データに属する雑音であるかを判別する雑音データ判別手段と、前記音入力手段から雑音の重畳された雑音重畳音声データに対して前記雑音データ判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記複数種類の音響モデルのうち前記雑音データ判別手段で判別された雑音データに対応した音響モデルを用いて音声認識を行う音声認識手段とを有したことを特徴としている。

【0029】

このように、本発明の音声認識装置は、現在の雑音が複数種類の雑音データのどの雑音データに属する雑音であるかを判別する雑音データ判別を行い、雑音重畳音声に対して、その雑音データ判定結果に基づいた雑音除去を行う。そして、その雑音除去音声に対し、当該雑音データに対応した音響モデルを用いて音声認識を行うようにしている。また、この音声認識装置が用いる複数種類の音響モデ

ルは、上述した音響モデル作成方法によって作成された音響モデルである。

【0030】

これによって、ある空間内に存在する雑音に対し最適な雑音除去処理が行えるとともに、その時点の雑音に対して最適な音響モデルを用いた音声認識が可能となるので、たとえば、自動車や作業場など特有の雑音環境下で高い認識性能を得ることができる。

【0031】

(7) 前記(6)の音声認識装置において、音入力手段に入力された雑音に対応する雑音パラメータを取得する雑音パラメータ取得手段を有するようにしている。

この雑音パラメータ取得手段を設けることによって、収集すべき雑音とその雑音の発生源との関連付けを確実に行うことができる。

【0032】

(8) 前記(6)または(7)の音声認識装置において、前記分類によって得られた複数種類の雑音データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うようにしている。

これによって、それぞれの雑音データに対して適正かつ効率的な雑音除去が可能となる。

【0033】

(9) 前記(6)から(8)のいずれかの音声認識装置において、雑音を有するある空間内は、乗り物内であることが1つの例として考えられる。

これによって、乗り物(たとえば、自動車)特有の様々な雑音の影響を考慮した音声認識を行うことができる。たとえば、運転者が乗り物自体または乗り物に搭載された機器などの操作や動作設定を行うような場合、高い認識精度で認識されることによって確実な操作や動作設定が可能となる。

【0034】

(10) 前記(9)の音声認識装置において、乗り物内で収集可能な各種雑音は、天候、前記乗り物の走行状況、前記乗り物の走行位置、前記乗り物に搭載されている機器の動作状態の少なくとも1つに起因する複数種類の雑音でとしてい

る。

これによって、乗り物（たとえば、自動車）特有の様々な雑音に対応可能な音響モデルを作成することができ、その音響モデルを用いることによって、その乗り物特有の様々な雑音の影響を考慮した音声認識が可能となり、高い認識精度を得ることができる。

【0035】

(11) 前記(6)から(10)のいずれかの音声認識装置において、音響モデルを作成するための前記雑音収集ステップは、前記収集すべき複数種類の雑音に対応するそれぞれの雑音パラメータを記録する工程を含み、前記雑音データ作成ステップは、前記収集すべき複数種類の雑音とこの収集すべき雑音に対応するそれぞれの雑音パラメータとを用いて分類することによって前記複数種類の雑音データを作成するようにしている。

【0036】

これによって、乗り物特有の雑音を適切に分類することができ、その分類によって得られた雑音データに対応した音響モデルを作成することができ、その音響モデルを用いることによって、その乗り物特有の様々な雑音の影響を考慮した音声認識が可能となり、高い認識精度を得ることができる。

【0037】

(12) 前記(6)から(11)のいずれかの音声認識装置において、前記複数種類の音響モデルを作成する際の雑音除去処理と前記認識すべき音声に対する音声認識を行う際の雑音除去処理は同じ雑音除去手法を用いるようにしている。

これによって、様々な雑音環境下において、高い認識精度を得ることができる。

【0038】

(13) 本発明の音声認識装置は、雑音を有するある空間内で音声認識を、前記(1)から(5)の音響モデル作成方法によって作成された複数種類の音響モデルを用いて行う音声認識装置であって、認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、前記音入力手段から入力された現在の雑音が、予め分類された複数種類の雑音データのどの雑音データに属する雑音であるかを判

別する雑音データ判別手段と、前記音入力手段から雑音の重畳された雑音重畳音声データに対して前記雑音データ判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記複数種類の音響モデルのうち前記雑音データ判別手段で判別された雑音種類に対応した音響モデルを用いて音声認識を行う音声認識手段とを有したことを特徴としている。

本発明の音声認識装置をこのような構成とすることによっても前記(6)の音声認識装置と同じ効果を得ることができる。

【0039】

(14) 本発明の音声認識装置を有する乗り物は、音声によって機器操作の可能な音声認識装置を有する乗り物であって、前記音声認識装置として、前記(6)から(13)のいずれかに記載の音声認識装置を有することを特徴としている。

これによって、たとえば、運転者が乗り物自体または乗り物に搭載された機器などの操作や動作設定を行うような場合、その乗り物特有の様々な雑音に適合した音響モデルを用いた音声認識が行えるので、高い認識精度を得ることができ、運転者などが音声で行う操作や動作設定が確実なものとなる。

【0040】

【発明の実施の形態】

以下、本発明の実施の形態について説明する。なお、この実施の形態で説明する内容は、本発明の音響モデル作成方法、音声認識装置、音声認識装置を備えた乗り物の説明を含むものである。

【0041】

また、本発明の実施の形態では、雑音を有する空間としては乗り物と工場を例にとって説明し、乗り物についての実施の形態を実施の形態1、工場についての実施の形態を実施の形態2として説明する。また、乗り物は自動車や2輪車などの車両の他、電車、飛行機、船舶など種々考えられるが、ここでは、自動車を例にとって説明する。

【0042】

〔実施の形態 1〕

まず、音声認識用の音響モデル作成方法の大まかな処理手順について図 1 のフローチャートを参照しながら簡単に説明する。これは、ここで説明する実施の形態 1 と後述する実施の形態 2 で共通である。

【0043】

まず、前記雑音を有する空間内で収集可能な各種雑音を収集する（ステップ S 1）。そして、これによって収集された雑音を分類して複数の雑音グループに対応した複数種類の雑音データを作成し（ステップ S 2）、この複数種類の雑音データを予め用意された標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成する（ステップ S 3）。続いて、この複数種類の雑音重畳音声データに対して雑音除去処理を行って複数種類の雑音除去音声データを作成し（ステップ S 4）、これによって作成された複数種類の雑音除去音声データから複数種類の音響モデルを作成する（ステップ S 5）。

【0044】

以下、自動車を例にとりて、本発明を詳細に説明する。まず、図 1 で説明した処理手順を図 2 を用いてより詳細に説明する。

自動車の場合、音声コマンド入力用のマイクロホンに入力される雑音の種類は自動車特有のものが多く、しかも、その雑音をあらかじめ収集しておくこともできる。

【0045】

そこで、自動車の室内で音声認識を行う際、音声認識性能に影響を与えそうな自動車特有の多種多様な雑音を収集し、収集された多種多様な雑音を統計的手法によって分類して、 n 個の雑音グループを作成し、それぞれの雑音グループごとの雑音データ N_1, N_2, \dots, N_n を作成する（これについて後に詳細に説明する）。

【0046】

なお、この n 個の雑音グループごとの雑音データ N_1, N_2, \dots, N_n (n 種類の雑音データ N_1, N_2, \dots, N_n) には、 S/N 比の違いも考慮されている。たとえば、同じ雑音でも S/N 比が 0 デシベルから 20 デシベル程度

に広がっている場合、 S/N 比の違いで雑音を分類して、 n 個の雑音グループに分けて、 n 種類の雑音データ N_1, N_2, \dots, N_n を作成する。

【0047】

そして、無響室などで収集された標準的な音声データ V （たとえば、多数の話者が多種類の単語について発話して得られた大量の音声データ）と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とを雑音重畳音声データ作成部 1 に与え、標準的な音声データ V と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とをそれぞれ重畳させて n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n を作成する。

【0048】

そして、雑音除去処理部 2 が n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対して最適な雑音除去処理手法を用いて雑音除去処理を行い、 n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を作成する。その後、音響モデル学習処理部 3 がこの n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を用いて音響モデルの学習を行い、 n 種類の音響モデル M_1, M_2, \dots, M_n を作成する。

【0049】

なお、 n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対するそれぞれ最適な雑音除去処理手法としては、 n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n ごとに用意された n 種類の雑音除去処理を用いて行うようにしてもよいが、数種類の代表的な雑音除去処理手法を用意して、その中からそれぞれの雑音重畳音声データにとって最適と思われる雑音除去処理手法を選択して用いるようにしてもよい。

【0050】

数種類の代表的な雑音除去処理手法としては、たとえば、前述したようなスペクトラム・サブストラクション (SS) やケプストラム平均化処理 (CMN)、さらには、音源を推定したエコーキャンセルなどがあり、これら雑音除去処理手法のうち、それぞれの雑音に最適な1つの雑音除去手法を選んで雑音除去を行うようにしてもよく、また、これらの雑音除去手法のうち2種類以上を組み合わせ

て、組み合わせたそれぞれの雑音除去手法に重み付けを行って雑音除去するようにしてもよい。

【0051】

次に、収集された多種多様な雑音を、ある統計的手法によって幾つか (n 個) に分類し、その分類によって得られたそれぞれの雑音グループごとの n 種類の雑音データ N_1, N_2, \dots, N_n を生成する具体的な例について図 3 を参照しながら詳細に説明する。

【0052】

この実施の形態 1 では、自動車に搭載された機器を操作するための音声コマンドを認識する際に本発明を適用する例であるので、雑音収集用の自動車を様々な条件で様々な状況のもとに長期間走行させて、自動車内の所定の場所に設置されたマイクロホン 11 から自動車特有の多種多様な雑音を時系列で収集する。

【0053】

なお、このマイクロホン 11 は、雑音収集用の自動車内において、運転者が音声によって機器操作を行う際、話者の音声コマンドを適正に入力できるような位置に設置することが望ましい。

【0054】

このマイクロホン 11 は、本発明の音声認識装置を実際に搭載するユーザ向け販売用の車種において、その設置位置がたとえばステアリング部分というように決められている場合にはその位置にマイクロホン 11 を設置して、そのマイクロホン 11 から雑音を収集する。そして、この収集された雑音は入力信号処理部 12 で増幅や A/D 変換などの入力信号処理が施されたのちに雑音記録部 22 に記録される。

【0055】

また、設計・開発段階などでマイクロホン 11 の設置位置が決定されていない場合は、設置候補となり得る複数の位置に設置して、それぞれのマイクロホン 11 から雑音を収集するようにしてもよい。この実施の形態では、設置位置はすでに決められているものとし、その設置位置に設置された 1 つのマイクロホン 11 から雑音を収集する例について説明する。

【0056】

また、マイクロホン11からの雑音収集と同時に、自動車の走行状況、現在位置、天候（ここでは雨量としている）、自動車に搭載されている各種機器の動作状態などを表す情報（雑音パラメータと呼ぶことにする）を時系列で収集する。

【0057】

この雑音パラメータは、自動車の速度を示す情報、エンジン回転数を示す情報、変速ギヤの位置を示す情報、窓の開閉状況（開度）を示す情報、エアコンの動作状態（風量の設定状態など）を示す情報、ワイパーの動作状態を示す情報、方向指示器の動作状態を示す情報、雨量計からの雨量を示す情報、GPS（Global Positioning System）による走行位置情報、カーオーディオの音信号を示す情報などであり、これら各雑音パラメータは、これらの雑音パラメータの取得が可能な雑音パラメータ取得部13から時系列に取得されて雑音パラメータ記録部21に記録される。

【0058】

なお、これら雑音パラメータ取得部13は自動車に設置されている。そして、たとえば、走行速度を示す情報を取得する速度情報取得部131、エンジン回転数を示す情報を取得する回転数情報取得部132、変速ギヤ位置を示す情報を取得する変速ギヤ位置情報取得部133、窓の開度をたとえば開度0%、開度50%、開度100%などというような情報として取得する窓開度情報取得部134、エアコンの動作状態をたとえば動作停止・風量（弱風・強風）などの情報として取得するエアコン動作情報取得部135、ワイパーのオン・オフ情報を取得するワイパー情報取得部136、方向指示器のオン・オフ情報を取得する方向指示器情報取得部137、GPSからの現在位置情報を取得する現在位置情報取得部138、雨量センサからの雨量情報（雨量なし、少量・多量など）を示す情報を取得する雨量情報取得部139、カーオーディオからの音量などの情報を取得するカーオーディオ情報取得部140などからなる。

【0059】

なお、上述したような実際に自動車を走行させてマイクロホン11から時系列に収集される雑音データと、雑音パラメータ取得部13の各情報取得部131～

140から時系列に取得される各雑音パラメータは、実際に自動車を走行（停止状態も含む）させることによって得られるものである。

【0060】

すなわち、その自動車をたとえば1ヶ月あるいは数ヶ月といった長期間、色々な場所、色々な天候のもとで走行させ、かつ、各雑音パラメータを色々変化させる。

たとえば、走行速度を色々変化させたり、エンジン回転数を色々変化させたり、変速ギヤを色々変化させたり、窓の開度を色々変化させたり、エアコンを色々な設定状態としたり、カーオーディオから色々な音信号を出力させたり、ワイパーや方向指示器などを適宜操作させるというように、自動車の走行時にあり得る様々な状態を作り出す。

【0061】

これによって、マイクロホン11からは多種多様な雑音が時系列で入力され、入力信号処理部12で増幅処理やデジタル信号への変換処理（A/D変換）がなされて、収集された雑音として雑音記録部22に記録されるとともに、その時々々の各雑音パラメータが雑音パラメータ取得部13で時系列に取得されて雑音パラメータ記録部21に記録される。

【0062】

そして、雑音分類処理部23が、マイクロホン11で収集された時系列の雑音（雑音記録部22に記録された時系列の雑音）と雑音パラメータ記録部21に記録された雑音パラメータとを用いて、ある統計的手法によってその収集された雑音を分類してn個の雑音グループを作成し、それぞれの雑音グループごとの雑音データ N_1 , N_2 , ..., N_n を生成する。

【0063】

この雑音分類処理部23が行う雑音分類には幾つかの手法が考えられるが、たとえば、収集された時系列の雑音データの特徴ベクトルをベクトル量子化して、そのベクトル量子化結果を用いてn個の雑音グループに分類する方法、あらかじめ用意された幾つかの音声認識用データに実際に重畳させてそれを実際に音声認識させて、その認識結果に基づいてn個の雑音グループに分類する方法などがある。

る。

【0064】

なお、 n 種類の雑音データ $N1, N2, \dots, Nn$ は、それぞれの雑音データ $N1, N2, \dots, Nn$ ごとに、走行速度を示す情報、回転数を示す情報、変速ギヤを示す情報、窓の開度を示す情報、エアコンの動作状態を示す情報など上述した様々な雑音パラメータの値に依存したものであるため、これらそれぞれの雑音パラメータと n 種類の雑音データ $N1, N2, \dots, Nn$ は互いに対応付けられたものとなっている。

【0065】

たとえば、雑音データ $N1$ は、走行速度が時速 40 km から時速 80 km の範囲内、回転数が 1500 rpm から 3000 rpm の範囲内、変速ギヤがトップギヤ、窓の開度は 0 (閉状態)、エアコンは弱風動作、ワイパーはオフ、 \dots (そのほかの雑音パラメータは省略) に対応した雑音データであって、雑音データ $N2$ は、走行速度が時速 80 km から 100 km の範囲内、回転数が 3000 rpm から 4000 rpm の範囲内、変速ギヤがトップギヤ、窓の開度は 50% (半開き状態)、エアコンは強動作、ワイパーはオフ、 \dots (そのほかの雑音パラメータは省略) に対応した雑音データであるといった具合である。

【0066】

これによって、現時点におけるそれぞれの雑音パラメータがそれぞれどのような値であるときは、そのときの雑音は n 種類の雑音データ $N1, N2, \dots, Nn$ のどの雑音データに属する雑音であるかを知ることができる。なお、 n 種類の雑音データ $N1, N2, \dots, Nn$ の具体例については後に説明する。

【0067】

このようにして、 n 種類の雑音データ $N1 \sim Nn$ が作成されると、図 2 に示すように、これら雑音データ $N1 \sim Nn$ を標準的な音声データ V (無響室などで収集された多数の話者が多種類単語について発話して得られた大量の音声データ) に重畳させ、 n 種類の雑音重畳音声データ $VN1, VN2, \dots, VNn$ を作成する。

【0068】

そして、この n 種類の雑音重畳音声データに対し、それぞれの雑音データ $N_1 \sim N_n$ を除去するに最適な雑音除去処理手法（前述したようにこの実施の形態 1 では、3 種類の雑音除去処理のどれか、またはそれらの組み合わせ）を用いて、雑音除去処理を行い、 n 個の雑音除去音声データ V_1' , V_2' , \dots , V_n' を作成し、この n 個の雑音除去音声データ V_1' , V_2' , \dots , V_n' を用いて音響モデルの学習を行い、 n 個の音響モデル M_1 , M_2 , \dots , M_n を作成する。

この n 個の音響モデル M_1 , M_2 , \dots , M_n は、 n 種類の雑音データ $N_1 \sim N_n$ に対応するものである。

【0069】

つまり、音響モデル M_1 は雑音データ N_1 の重畳した音声データ（雑音重畳音声データ V_{N1} ）から雑音データ N_1 を除去した後（雑音データ N_1 は完全には除去されないでその成分が残っている）の音声データ V_1' から作られた音響モデルであり、音響モデル M_2 は雑音データ N_2 の重畳した音声データから雑音データ N_2 を除去した後（雑音データ N_2 は完全には除去されないでその成分が残っている）の音声データから作られた音響モデルである。

【0070】

また、音響モデル M_n は雑音データ N_n の重畳した音声データ（雑音重畳音声データ V_{Nn} ）から雑音データ N_n を除去した後（雑音データ N_n は完全には除去されないでその成分が残っている）の音声データ V_n' から作られた音響モデルである。

以上のようにして、本発明の実施の形態 1 である自動車の機器操作を音声によって行う際の音声認識に用いられる音響モデル M_1 , M_2 , \dots , M_n が作成される。

【0071】

次に、このような音響モデル M_1 , M_2 , \dots , M_n を作成する際の雑音データ（マイクロホン 11 から収集された雑音）の分類処理について具体的に説明する。

【0072】

雑音収集を行うために長期間自動車を走行させることによって収集された雑音には様々な雑音が含まれる。たとえば、タイヤのパターンノイズ（主に速度に係する）やエンジン音（主に速度、エンジンの回転数、ギヤ位置に係する）、窓が開いている場合の風切り音、エアコンの動作音、雨が降っていれば雨そのものの音やワイパーの動作音、方向変換時には方向指示器の動作音、トンネル通過時は反響音、カーオーディオの動作中には音楽などの音信号などが収集される。

【0073】

そして、ある時刻においては、これらがすべて雑音として収集される場合もあり、また、ある時刻においては、これらのうち、たとえば、タイヤのパターンノイズやエンジン音だけが収集される場合もある。また、このような雑音とともに、それぞれの時刻対応にその自動車に設置された各種の雑音パラメータ取得部 13 で取得された雑音パラメータが記録されている。

【0074】

本来、雑音は前述したように多種存在し、マイクロホン 11 からは個々の雑音パラメータに対応した雑音、また、複数の雑音パラメータの組み合わせに対応した数多くの種類の雑音が収集され、このマイクロホン 11 から収集された雑音を統計的手法によって実用的な数の雑音グループ数とするための分類処理がなされる。しかし、ここでは、説明をより簡素化するために、3 種類の雑音パラメータ（走行速度、エアコンの動作状態、雨量）のみで考え、これら走行速度、エアコンの動作状態、雨量の 3 つの雑音パラメータを 3 次元座標における 3 つの直交軸上の値（ここではそれぞれ 3 段階の状態を示す値）で表して分類する例について説明する。

【0075】

この場合、速度は「停車時（速度 0）」、「低速」、「高速」の 3 段階で表し、エアコンの動作状態は、「停止」、「弱風」、「強風」の 3 段階で表し、雨量は「無」、「少量」、「多量」の 3 段階で表すものとする。

【0076】

なお、速度の「低速」および「高速」は、たとえば、60 km/h までを低速、それ以上を高速というように予め範囲を決めておく。また、雨量も同様に、雨

量計から得られる 1 時間雨量が 0 mm を「無」、同じく、雨量計から得られる 1 時間雨量が 5 mm までを「少量」、それ以上を「多量」というように予め範囲を決めておく。

【0077】

また、この雨量を示す雑音パラメータ（「無」、「少量」、「多量」）は、雨量計からではなく、ワイパーの動作状態を用いることもできる。たとえば、ワイパーがオフしていれば、雨量は「無」、ワイパーが低速動作であれば雨量は「少量」、ワイパーが高速動作であれば雨量は「多量」というように判定できる。

【0078】

図 4 は上述の 3 種類の雑音パラメータによる雑音を収集対象とし、これら 3 種類の雑音パラメータに対応して発生する雑音を 1 つのマイクロホン 11 を用いて長期間かけて収集した雑音データ（これを N で表す）を 1 つの大きな球体で表している。この図 4 は速度を「停車時」、「低速」、「高速」の 3 段階、エアコンの動作状態を「停止」、「弱風」、「強風」の 3 段階、雨量を「無」、「少量」、「多量」の 3 段階として、それらを 3 次元座標上で表したものである。

【0079】

この雑音データ N に対し、ベクトル量子化などを用いた統計的手法を用いることなく単純に、個々の雑音パラメータごとに分類すると、図 5 のようになる。この場合、3 の 3 乗個（27 個）の雑音グループが得られ、それぞれの雑音グループに対応した 27 個の雑音データ N1 ～ N27 が得られる。この 27 個の雑音データ N1 ～ N27 をそれぞれ小さな球体で表している。

【0080】

この図 5 において、幾つかの雑音データについて説明すると、たとえば、雑音データ N1 は速度が「停車時（速度 0）」、エアコンは「停止」、雨量は「無」に対応する雑音データであり、雑音データ N5 は速度が「低速」、エアコンは「弱風」、雨量は「無」に対応する雑音データであり、雑音データ N27 は速度が「高速」、エアコンは「強風」、雨量は「多量」に対応する雑音データである。

【0081】

なお、この図 5 では、個々の雑音データ N1 ～ N27 を雨量が「無」、「少量

」、「多量」で色の濃さを分けして表しており、雨量が「無」における 3×3 個の雑音データ $N1 \sim N9$ を最も薄い色で表し、雨量が「少量」における 3×3 個の雑音データ $N10 \sim N18$ を中程度の濃さで表し、雨量の「多量」における 3×3 個の雑音データ $N19 \sim N27$ を最も濃い色で表している。

【0082】

この図5によれば、確かに、自動車の現時点における雑音パラメータによって、マイクロホン11にはどのような種類の雑音データが入力されるかを知ることができ、それによって、最適な音響モデルを用いて音声認識することが可能となる。たとえば、現在の自動車の速度が「低速」でエアコンが「弱風」で雨量「無」であれば、そのときの雑音データは $N5$ であり、その雑音データ $N5$ に対応した音響モデルを用いて音声認識を行う。

【0083】

この図5の場合は、マイクロホン11から得られた時系列の雑音データを単純に個々の雑音パラメータのとり得る状況の数（この例では27種類）ごとに分類した例であるが、これをある統計的手法によって分類した例について図6により説明する。

【0084】

なお、このようなある統計的手法を用いて分類する例として、前述したように、雑音データの各時刻対応の特徴ベクトルをベクトル量子化して、そのベクトル量子化結果を用いて、複数個の雑音グループに分類する方法、あらかじめ用意された幾つかの音声認識用データに実際に重畳させてそれを実際に音声認識させて、その認識結果に基づいて複数個の雑音グループに分類する方法などがある。

【0085】

このような手法によって分類された結果、図6に示すように9個の雑音グループが作成され、それぞれの雑音グループに対応した9種類の雑音データ $N1 \sim N9$ が作成されたとする。

この図6の場合、雨の音（雨量）が音声認識する上での雑音データとして最も影響が大きく、その次に自動車の走行速度の影響が大きく、エアコンの影響は雨や走行速度に比べて影響が小さいことを示している。

【0086】

この図6では、雨量が「無」の場合、自動車の走行速度が0（「停車時」）では、エアコンの動作状態に対応した雑音データN1, N2, N3が作成されている。また、自動車の走行速度が「低速」においては、エアコンの動作状態が「停止」に対応した雑音データN4とエアコンの動作状態が「弱風」・「強風」で1つの雑音データN5が作成される。すなわち、自動車がある速度で走行している場合には、エアコンの動作状態が「弱風」であっても「強風」であってもその動作音は、自動車の走行による雑音に比べるとその影響は殆どないと判断された結果である。また、自動車の速度が「高速」においては、エアコンの動作状態に関係なく、1つの雑音データN6が作成されている。

【0087】

また、雨が降っていると、たとえ、雨量が「少量」であっても、エアコンがどのような動作状態にあるかは関係なく、自動車の走行速度に依存した雑音データが作成される。すなわち、雨量が「少量」である場合には、走行速度が「低速」まで（停車時を含む）に対応した雑音データN7と、「高速」に対応した雑音データN8の2種類の雑音グループが作成されている。

また、雨が「多量」である場合には、エアコンの動作状態や自動車の走行速度の影響は殆どなく、1つの雑音データN9が作成されている。

【0088】

以上のように、3種類の雑音パラメータ（走行速度、エアコンの動作状態、雨量）に対応する雑音を収集対象とし、これら3種類の雑音パラメータに依存する雑音を1つのマイクロホン11を用いて長期間かけて収集した雑音データNを、ある統計的手法で分類した結果、図6のような雑音データN1～N9が作成されたとする。

【0089】

なお、この図6により得られた雑音データN1～N9は、説明をわかり易くするため、雑音パラメータを3つ（走行速度、エアコンの動作状態、雨量）とした例であるが、実際には、雑音パラメータは、前述したように、多種類存在し、これら多種類の雑音パラメータに依存した各種雑音を長期間収集して、時系列の雑

音データを得て、その時系列の雑音データを統計的手法によって分類し、 n 個の雑音グループを得て、これらそれぞれの雑音グループに対応する n 種類の雑音データ $N_1 \sim N_n$ を作成する。

【0090】

また、実用的な雑音グループ数としては、音響モデル作成処理や音声認識処理の効率化などの点から数個から 10 数個程度が好ましいが、これは、任意に設定することができる。

【0091】

このようにして、 n 個の雑音グループに対応する n 種類の雑音データ $N_1 \sim N_n$ が作成されると、前述したように（図 1 参照）、この n 種類の雑音データ $N_1 \sim N_n$ を標準的な音声データに重畳させて n 個の雑音重畳音声データ VN_1, VN_2, \dots, VN_n を作成し、この n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対し、それぞれの雑音を除去するに最適な雑音除去処理手法を用いて、雑音除去処理を行い、 n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を作成する。

【0092】

そして、この n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を用いて音響モデルの学習を行い、 n 種類の音響モデル M_1, M_2, \dots, M_n を作成する。これによって、 n 種類の雑音データ N_1, N_2, \dots, N_n に対応した n 種類の音響モデル M_1, M_2, \dots, M_n を作成することができる。

【0093】

次に上述のようにして作成された n 種類の音響モデル M_1, M_2, \dots, M_n を用いた音声認識について説明する。

【0094】

図 7 は本発明の音声認識装置の構成図であり、機器操作の音声コマンドや様々な雑音を入力する音入力手段としてのマイクロホン 11、このマイクロホン 11 から入力された音声コマンドを増幅するとともにデジタル信号に変換（A/D 変換）する入力信号処理部 12、前述した各種雑音パラメータを取得する雑音パラメータ取得部 13、この雑音パラメータ取得部 13 から取得された各種の雑

音パラメータに基づいて、現在の雑音の種類が前述した n 個に分類されることによって作成された n 種類の雑音データ $N1 \sim Nn$ のどれに相当するかを判別する雑音データ判定部 14、雑音データ $N1 \sim Nn$ それぞれに最適な雑音除去方法が保存された雑音除去方法保存部 15、雑音データ判定部 14 で判別された雑音データに最適な雑音除去方法を、雑音除去方法保存部 15 に保存された各種の雑音除去方法の中から選択し、マイクロホン 11 から入力された音声データ（デジタル変換後の雑音重畳音声データ）に対して雑音除去処理を行う雑音除去処理部 16、この雑音除去処理部 16 によって雑音除去された雑音除去音声データに対し、前述の手法によって作成された音響モデル $M1 \sim Mn$ （ n 種類の雑音データ $N1 \sim Nn$ に対応）のうちのいずれかの音響モデルと言語モデル 17 を用いて音声認識を行う音声認識処理部 18 を有した構成となっている。

【0095】

この図 7 で示される音声認識装置は、乗り物（この実施の形態では自動車）内の適当な場所に設置される。

図 8 は図 7 で示される音声認識装置（図 8 においては、30 の符号が付されている）が設置された乗り物（この図 8 の例では自動車）の一例を示すもので、音声認識装置 30 は、自動車室内の適当なスペースに取り付けられる。なお、音声認識装置 30 の設置位置はこの図 8 の例に限られるものではなく、たとえば、座席と床の間のスペースや荷物室など適当な場所を選ぶことができることは勿論である。また、音声認識装置 30 のマイクロホン 11 は、運転者が音声を入力しやすい位置として、たとえば、ステアリング 31 部分に設けられるが、これもステアリング 31 部分に限られるものではない。

【0096】

ところで、図 7 に示す雑音データ判定部 14 は、雑音パラメータ取得部 13 から各種雑音パラメータを受け取って、マイクロホン 11 から入力された現在の雑音が、複数種類の雑音データ $N1 \sim N9$ のどの雑音データに属するかを判定するものである。

【0097】

すなわち、この雑音データ判定部 14 は、雑音パラメータ取得部 13 からの雑

音パラメータとして、たとえば、前述したように、速度情報取得部 131 からの速度を示す情報、エアコン動作情報取得部 135 からのエアコン動作状態を示す情報、雨量情報取得部 139 からの雨量を示す情報などに基づいて、現在の雑音状況が雑音データ N1～N9 のどの雑音データに属するかを判定する。

【0098】

たとえば、現在の走行速度が 70 km、エアコンの動作状態が「弱風」、雨量が「無」といった情報を雑音パラメータとして雑音データ判定部 14 が受け取ると、これらの雑音パラメータから、現在の雑音が雑音データ N1～N9 のどの雑音データに属するかを判定する。仮に、現在の雑音が雑音データ N6 に属すると判定されたとすると、その判定結果を雑音除去処理部 16 と音声認識処理部 18 に送る。

【0099】

雑音除去処理部 16 は、雑音データ判定部 14 からの現在の雑音種類を示す情報を受け取ると、入力信号処理部 12 からの雑音重畳音声データに対して最適な雑音除去方法を用いた雑音除去処理を行う。たとえば、雑音データ判定部 14 から現在の雑音が雑音データ N6 に属する雑音であることを示す情報が雑音除去処理部 16 に与えられると、この雑音除去処理部 16 は、雑音データ N6 に最適な雑音除去手法を雑音除去方法保存部 15 から選択し、選択した雑音除去方法によって雑音重畳音声データに対する雑音除去処理を行う。

【0100】

なお、この雑音除去処理は、この実施の形態の場合、前述したように、たとえば、スペクトラム・サブトラクション (SS) やケプストラム平均化処理 (CMN) などのいずれかあるいはそれらの組み合わせのどれかによって行う。

【0101】

また、現在の雑音の中にカーオーディオからの音信号やワイパーの動作音、方向指示器の動作音が含まれる場合は、これらの雑音を直接除去するような処理も可能とする。

【0102】

たとえば、マイクロホン 11 に入力された雑音重畳音声データに含まれるカー

オーディオからの音信号に対しては、カーオーディオから直接得られる音信号、すなわち、カーオーディオ情報取得部 140 から得られるカーオーディオ信号を雑音除去処理部 16 に与え（図 7 において一点鎖線で示す）、そのカーオーディオ信号をマイクロホンに入力された雑音重畳音声データから差し引くことで、マイクロホン 11 に入力された雑音重畳データに含まれるカーオーディオの音信号成分を除去することができる。このとき、雑音除去処理部 16 では、マイクロホン 11 からの雑音重畳音声データに含まれるカーオーディオ信号は、カーオーディオから直接得られる信号に比べて一定の時間遅れが生じるので、その時間遅れ分を考慮した除去処理を行う。

【0103】

また、ワイパーや方向指示器の動作音は、周期的な動作音であり、それぞれの周期や雑音成分（動作音）は車種によって決められているので、その周期に対応したタイミング信号（図 7 において一点鎖線で示す）をワイパー情報取得部 136 や方向指示器情報取得部 137 から雑音除去処理部 16 に送ることで、雑音除去処理部 16 では、そのタイミングでワイパーの動作音や方向指示器の動作音を除去することができる。この場合も、マイクロホン 11 からの雑音重畳音声データに含まれるワイパーの動作音や方向指示器の動作音は、ワイパーや方向指示器から直接得られる動作信号に比べて一定の時間遅れが生じるので、その時間遅れ分を考慮したタイミングで雑音除去処理を行う。

【0104】

以上のようにして、マイクロホン 11 から入力されたある時刻の雑音重畳音声データ（音声コマンドとその時点でマイクロホンに入力される雑音からなる）に対する雑音除去処理がなされると、その雑音の除去された雑音除去音声データは音声認識処理部 18 に送られる。

【0105】

この音声認識処理部 18 にも、雑音データ判定部 14 から雑音データ判定結果として、雑音データ N1～N9 のいずれかを示す情報が与えられており、その雑音データ判定結果に対応する音響モデルを選択し、選択した音響モデルと言語モデル 17 を用いて音声認識処理を行う。たとえば、マイクロホン 11 に入力され

た話者からの音声コマンドに重畳された雑音が雑音データ N 1 に属する雑音であることを示す情報を雑音データ判定部 14 から受け取ったとすると、音声認識処理部 18 では、音響モデルとしては、雑音データ N 1 に対応する音響モデル M 1 を選択する。

【0106】

この音響モデル M 1 は、前述の音響モデル作成方法で説明したように、雑音データ N 1 を音声データに重畳させ、その雑音重畳音声データから雑音除去して雑音除去音声データを作成し、その雑音除去音声データから作成された音響モデルであるため、話者の発した音声コマンドに重畳された雑音が雑音データ N 1 に属する場合には、その音声コマンドにとって最適な音響モデルとなり、認識性能を高めることができる。

【0107】

1つの具体例として、図6に示すような9個の雑音グループに対応する9種類の雑音データ N 1 ～N 9 が作成され、これら9種類の雑音データ N 1 ～N 9 に対応した音響モデル M 1 ～M 9 が作成されている場合の音声認識動作について説明する。

【0108】

今、運転者が動作中に音声コマンドを与えると、音声認識装置 30 側ではその音声コマンドを認識して、その認識結果に基づいた機器操作を行う場合を考える。なお、このとき自動車の走行速度が時速 40 km（低速走行であるとする）、エアコンの動作状態は「弱風」、雨量は「無」とする。

【0109】

この場合、自動車内のある位置（ステアリングなど）に設置されたマイクロホン 11 には、そのときの状況に応じた雑音が入力され、その状態で運転者が何らかの音声コマンドを発すると、その音声コマンドにはそのときの状況に応じた雑音が重畳され、その雑音重畳音声データは入力信号処理部 12 で増幅や A/D 変換されたのちに雑音除去処理部 16 に送られる。

【0110】

一方、雑音データ判定部 14 には、この場合、現在の雑音パラメータとして、

雑音パラメータ取得部 13 の速度情報取得部 131 から現時点の走行速度を示す情報、エアコン動作情報取得部 135 からエアコンの動作状態を示す情報、雨量情報取得部 139 からの雨量を示す情報が雑音パラメータとして与えられており、それらの雑音パラメータに基づいて、現在の雑音がどの雑音データ N1 ~ N9 のうちのどの雑音データに属する雑音であるかを判定する。

【0111】

この場合、走行速度を示す情報は時速 40 km（ここでは「低速」としている）、エアコンの動作状態を示す情報は「弱風」、雨量を示す情報は「無」であるので、雑音データ判定部 14 は、図 6 から現在の雑音は雑音データ N5 であると判定し、その判定結果を雑音除去処理部 16 と音声認識処理部 18 に送る。

【0112】

これによって、雑音除去処理部 16 では、その雑音データ N5 に最適な雑音除去処理手法を用いて雑音除去処理を行い、その雑音除去音声データを音声認識処理部 18 に送る。

【0113】

音声認識処理部 18 では、雑音データ判定部 14 から送られてきた雑音データ N5 に対応した音響モデル M5（図 7 では図示されていない）を選択し、その音響モデル M5 と言語モデル 17 を用いて、雑音除去処理部 16 で雑音除去された雑音除去音声データに対して音声認識処理を行う。そして、この音声認識結果に基づいて機器操作を行う。この機器操作の例としては、たとえば、ナビゲーションシステムに対する目的地の設定などである。

【0114】

以上のように、この実施の形態 1 における音声認識装置では、音声コマンドに重畳された雑音が雑音データ N1 ~ N9 のいずれかに属するかを判定し、それに対応した雑音除去処理手法（音響モデル作成時と同じ雑音除去処理手法）を用いて雑音除去を行い、その雑音除去のなされた音声データ（雑音除去音声データ）に対し、最適な音響モデルを用いて音声認識を行うようにしている。

【0115】

すなわち、自動車のその時々々の走行状況、走行位置、車載機器の動作状態など

に対応した様々な種類の雑音が音声コマンドに重畳されても、それに対応した最適な雑音除去が行え、その雑音除去後の音声データに対し、最適な音響モデルを用いて音声認識することができるので、様々な雑音環境下において高い認識性能を得ることができる。

【0116】

特に、自動車において、車種が限定されているような場合に特に有効なものとなる。すなわち、雑音収集を行って音響モデルを作成するための雑音収集用車種と実際に本発明の音声認識装置を搭載するユーザ向け販売用車種が同じであれば、雑音収集用車種における雑音収集のためのマイクロホン取り付け位置とユーザ向け販売用車種における音声コマンド入力用のマイクロホン取り付け位置を同じとすることで、雑音がほぼ同じ条件でマイクロホンから入力されるので、適正な音響モデルを選択することができ、高い認識性能を得ることができる。

【0117】

なお、音響モデルを作成するための雑音収集用自動車は専用を用意することも可能であるが、ユーザ向け販売用自動車に、音響モデル作成（図3で示した雑音データN1～Nnの作成を含む）を行うための機能を音声認識装置30とともに搭載して、1台の自動車で音響モデル作成機能と音声認識機能の両方を可能とすることも可能である。その場合、マイクロホン11、入力信号処理部12、雑音パラメータ取得部13、雑音除去処理部16などは音響モデル作成時と音声認識時で共用することができる。

【0118】

このように、ユーザ向け販売用自動車に音響モデル作成機能と音声認識機能の両方を持たせることで、雑音環境の変化などにより雑音の分類の変更が容易に行え、それに伴って、音響モデルを新たに生成したり、更新したりすることが可能となり、雑音環境の変化への対応がし易くなる。

【0119】

〔実施の形態2〕

この実施の形態2では、雑音を有する空間としては工場の作業場を例にとって説明する。たとえば、ベルトコンベアによって搬送されてくる物品などの検査結

果の記録など音声として入力し、それを音声認識してその認識結果を検査記録として保存するような状況を考える。

【0120】

図9は工場内のある作業場を示すもので、作業場41内には、製品を加工する加工装置42、この加工装置42によって加工された製品を搬送するベルトコンベア43、ベルトコンベア43によって搬送された製品を検査する検査装置44、作業場41内の温度や湿度などを調整するエアコンディショナ（エアコン）45、作業者（図示せず）の発話する音声を認識する本発明の音声認識装置30などが図9のように設置されているとする。

【0121】

また、図示のP1、P2、P3は作業者（図示されていない）が何らかの作業を行って、その位置で音声入力を行う位置とする。すなわち、作業者は位置P1で何らかの作業を行い、そのあと、位置P2に移動して何らかの作業を行い、さらに、位置P3に移動して検査装置44で検査を行うといった行動をとるものとし、図示の太線Aはその動作線（以下では動作線Aという）を示している。

【0122】

そして、加工装置42から送出される製品について、作業者は位置P1、P2において、それぞれの位置P1、P2における確認項目に対する確認結果などを音声によって入力し、位置P3では検査装置44を用いた検査を行って、その検査結果を音声で入力するといった作業を行うものとする。

【0123】

なお、作業者はヘッドセット型のマイクロホンを着用し、そのマイクロホンから入力された音声は音声認識装置30に送信されるようになっている。そして、音声認識装置30で音声認識されたそれぞれの位置P1、P2、P3における確認結果や検査結果は、図9では図示されない記録手段に記録されるようになっている。

【0124】

ところで、このような作業場41での音声認識を行うには、作業場41特有の雑音を考慮する必要があるが、前述の実施の形態1で説明した自動車と同様、そ

の雑音をあらかじめ収集しておくことができる。

【0125】

そこで、このような作業場 41 で音声認識を行う際、音声認識性能に影響を与えそうな作業場 41 特有の多種多様な雑音を収集し、前述の実施の形態 1 で図 2 を用いて説明したと同様に、収集された多種多様な雑音を分類して、 n 個の雑音グループを作成し、それぞれの雑音グループごとの雑音データ N_1, N_2, \dots, N_n (n 種類の雑音データ N_1, N_2, \dots, N_n) を作成する。

【0126】

そして、無響室などで収集された標準的な音声データ V (たとえば、多数の話者が多種類の単語について発話して得られた大量の音声データ) と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とを雑音重畳音声データ作成部 1 に与え、標準的な音声データ V と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とをそれぞれ重畳させて n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n を作成する。

【0127】

そして、雑音除去処理部 2 が n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対して最適な雑音除去処理手法を用いて雑音除去処理を行い、 n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を作成する。その後、音響モデル学習処理部 3 がこの n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を用いて音響モデルの学習を行い、 n 種類の音響モデル M_1, M_2, \dots, M_n を作成する。

【0128】

なお、 n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対するそれぞれ最適な雑音除去処理手法については、前述の実施の形態 1 で説明したと同様に考えることができる。

【0129】

次に、収集された多種多様な雑音を、 n 個に分類し、分類されたそれぞれの雑音グループごとの雑音データ N_1, N_2, \dots, N_n を生成する具体的な例について図 10 を参照しながら詳細に説明する。

【0130】

この実施の形態2では、その作業場41で通常用いる加工装置42、ベルトコンベア43、検査装置44、エアコン45などを通常の作業時と同様の動作状態として、所定の期間、雑音収集を行う。この雑音収集は、作業者がたとえばヘッドセットなどを装着し、そのヘッドセットに設けられたマイクロホン11から、その作業場特有の多種多様な雑音データをある期間、時系列で収集する。

なお、このとき、作業者は自分が行う実際の作業を行いながらヘッドセットに設けられたマイクロホン11から様々な雑音を入力する。

【0131】

この実施の形態2では、作業者は図9に示すような作業場41において、動作線Aに沿って移動しながら作業を行うとしているので、作業者の移動に伴い、その動作線A上の作業者の位置を入力しながら雑音収集を行う。なお、作業者が、予め決められた位置のみで作業を行う場合には、その位置にマイクロホン11を設置して雑音収集を行うことができる。

【0132】

また、マイクロホン11からの雑音収集と同時に、作業場41での雑音の発生源となる機器の動作状態などを表す情報としての雑音パラメータを雑音パラメータ取得部13で時系列に取得する。

【0133】

この実施の形態2の場合、取得する雑音パラメータは、加工装置42の動作状態を示す情報（動作速度とする）、エアコン45の動作状態を示す情報（風量とする）、ベルトコンベア43の動作状態を示す情報（動作速度とする）、検査装置44の動作状態を示す情報（たとえば、この検査装置44による検査方式が複数種類あって、その種類によって検査装置44の発生する音が異なる場合、その検査方式の種類を表す情報とする）、作業者の位置（たとえば、作業者の図9で示した動作線A上の1次元座標、または、作業場41の床上の2次元座標、または、図9で示したような位置P1、P2、P3というような離散値）、作業場に設けられている窓や扉の開閉状況（窓や扉の開度とする）、作業場内に流れる放送などの有無やその内容、周囲の荷物の状況など様々である。

【0134】

なお、雑音パラメータ取得部13は作業場41に設置され、上述したような様々な雑音パラメータを取得するために、たとえば、加工装置42がどのような速度で動作しているかを示す情報を取得する加工装置動作情報取得部151、エアコン45がどのような動作状態となっているかを示す動作情報を取得するエアコン動作情報取得部152、ベルトコンベア43がどのような速度で動作しているかを示すベルトコンベア動作情報取得部153、検査装置44の動作情報を取得する検査装置動作情報取得部154、作業者が現在どの位置にいるかの位置情報を取得する作業者位置情報取得部155、窓の開度を示す情報を取得する窓開度情報取得部156などからなる。なお、取得すべき雑音パラメータはこれ以外に種々考えられるがそれらの図示は省略する。

【0135】

なお、マイクロホン11から時系列に収集される雑音と、雑音パラメータ取得部13の各情報取得部151～156から時系列に取得される各雑音パラメータは、作業者が実際に作業場41で作業を行うことによって得ることができるものである。

【0136】

すなわち、作業場41でたとえば1ヶ月といった期間、その作業場で発生し得る雑音を得るために、加工装置42、ベルトコンベア43、検査装置44、エアコン45などの機器の動作状態を変えたり、窓の開度を色々変化させたりというように、作業場であり得る様々な雑音環境を作り出す。

【0137】

これによって、マイクロホン11からは多種多様な雑音が時系列で入力され、入力信号処理部12で増幅処理やデジタル信号への変換処理(A/D変換)がなされて、収集された雑音として雑音記録部22に記録されるとともに、その時々各種雑音パラメータが雑音パラメータ取得部13で時系列に取得されて雑音パラメータ記録部21に記録される。

【0138】

そして、雑音分類処理部23が、マイクロホン11で収集された時系列の雑音

(雑音記録部 22 に記録された時系列の雑音) と雑音パラメータ記録部 21 に記録された雑音パラメータとを用いて、ある統計的手法によってその収集された雑音を分類して n 個の雑音グループを作成し、それぞれの雑音グループごとの雑音データ N_1, N_2, \dots, N_n を生成する。

【0139】

本来、雑音は前述したように多種存在し、マイクロホン 11 からは個々の雑音パラメータに対応した雑音、また、複数の雑音パラメータの組み合わせに対応した数多くの種類の雑音が収集され、このマイクロホン 11 から収集された雑音を統計的手法によって実用的な数の雑音グループ数とするための分類処理がなされる。しかし、ここでは、説明を簡素化するために、雑音パラメータを 3 種類の雑音パラメータ (作業者の位置、加工装置 42 の動作状態、エアコン 45 の動作状態) のみで考え、これら作業者の位置、加工装置の動作状態、エアコン 45 の動作状態の 3 つの雑音パラメータを、3 次元座標における 3 つの直交軸上の値 (ここではそれぞれ 3 段階の状態を示す値) で表して分類する例について説明する。

【0140】

すなわち、作業者の位置を図 9 における位置 P_1, P_2, P_3 の 3 つの位置で表し、加工装置 42 の動作状態は、この場合、「停止」、「低速」、「高速」の 3 段階で表し、エアコンの動作状態は、「停止」、「弱風」、「強風」の 3 段階で表すものとする。

【0141】

図 11 は上述の 3 種類の雑音パラメータに対応した雑音を前述の実施の形態 1 で説明したと同様の分類処理 (実施の形態 1 の説明で用いた図 4 の状態から図 5 の状態となるような分類処理) を行い、それをさらに、ある統計的手法による分類処理 (実施の形態 1 の説明で用いた図 5 の状態から図 6 の状態となるような分類処理と同様の分類処理) を行うことによって得られた分類結果の一例である。

この図 11 では、それぞれの雑音グループに対応した 12 種類の雑音データ $N_1 \sim N_{12}$ が 3 次元座標上に示されている。この 3 次元座標上の 12 種類の雑音データ $N_1 \sim N_{12}$ を、加工装置の 3 つの動作状態「停止」、「低速」、「高速」それぞれにおける 2 次元断面で表したものが図 12 (a) ~ (c) である。

【0142】

図12(a)は加工装置42が「停止」の場合であり、この場合は、作業者の位置P1, P2, P3に応じて、エアコン45の影響を受けた雑音データN1, N2, N3, N4, N5, N6が作成される。

【0143】

すなわち、作業者の位置がエアコン45から遠い位置P1では、エアコン45の動作状態（「停止」、「弱風」、「強風」）に関係しない1つの雑音データN1が作成され、作業者の位置がP2では、エアコン45の動作状態が「停止」か否かによって、それぞれに対応した雑音データN2, N3が作成される。なお、「停止」の場合は雑音データN2、「弱風」および「強風」のいずれの場合も1つの雑音データN3が作成される。

【0144】

また、作業者の位置がP3では、エアコン45の動作状態が「停止」の場合は雑音データN4が作成され、エアコン45の動作状態が「弱風」の場合は雑音データN5が作成され、エアコン45の動作状態が「強風」の場合は雑音データN6が作成されるというように、エアコン45の動作状態それぞれに対応する雑音データが作成される。

【0145】

これは、加工装置42の動作が停止しているときは、作業者の位置P1, P2, P3における雑音は、エアコン45の動作状態の影響を大きく受け、かつ、位置P1, P2, P3によってその影響の受け方が異なっていることを示している。

【0146】

また、図12(b)は加工装置42が「低速」の場合であり、この場合は、作業者の位置P1, P2, P3に応じて、加工装置42の影響が反映された雑音データN7, N8, N9, N10が作成される。

【0147】

すなわち、作業者が位置P1では、エアコン45の動作状態（「停止」、「弱風」、「強風」）に関係しない雑音データN7が作成され、作業者が位置P2では

、同じくエアコン45の動作状態（「停止」、「弱風」、「強風」）に関係しない雑音データN8が作成される。また、作業者の位置がP3では、エアコン45の動作状態が「停止」の場合は雑音データN9が作成され、エアコンの動作状態が「弱風」と「強風」で1つの雑音データN10が作成される。

【0148】

また、図12（c）は加工装置42の動作状態が「高速」の場合であり、この場合は、加工装置42の影響を大きく受けた雑音データN11、N12が作成される。

【0149】

すなわち、作業者の位置がP1、P2のいずれであっても、エアコン45の動作状態（「停止」、「弱風」、「強風」）に関係しない1つの雑音データN11が作成される。また、作業者の位置がエアコン45に近い位置P3では、エアコン45の影響も多少は反映されるが、エアコン45の動作状態（「停止」、「弱風」、「強風」）に関係しない1つの雑音データN12が作成される。

【0150】

この図12からもわかるように、加工装置42の動作が停止しているときは、作業者の位置P1、P2、P3における雑音は、位置P1、P2、P3それぞれによってエアコン45の動作音の影響を大きく受け、加工装置42の動作中は、位置によってはエアコン45の影響も多少反映されるが、加工装置42の動作音が全体の雑音を支配する傾向にあることを示している。

【0151】

以上のように、3種類の雑音パラメータ（作業者の位置、加工装置42の動作状態、エアコン45の動作状態）に依存する雑音を、マイクロホン11を用いて長期間かけて収集し、その収集した雑音を、ある統計的手法で分類した結果、図11のような雑音データN1～N12が作成されたとする。

【0152】

このように、n個（この例では12個）の雑音グループに対応する12種類の雑音データN1～N12が作成されると、図1で説明したように、この12種類の雑音データN1～N12を標準的な音声データに重畳させて12個の雑音重畳

音声データ $VN1$, $VN2$, \dots , $VN12$ を作成する。そして、この 12 種類の雑音重畳音声データ $VN1$, $VN2$, \dots , $VN12$ に対し、それぞれの雑音を除去するに最適な雑音除去処理手法を用いて、雑音除去処理を行い、12 種類の雑音除去音声データ $V1'$, $V2'$, \dots , $V12'$ を作成する。

【0153】

そして、この 12 種類の雑音除去音声データ $V1'$, $V2'$, \dots , $V12'$ を用いて音響モデルの学習を行い、12 種類の音響モデル $M1$, $M2$, \dots , $M12$ を作成する。

これによって、12 種類の雑音データ $N1$, $N2$, \dots , $N12$ に対応した 12 種類の音響モデル $M1$, $M2$, \dots , $M12$ を作成することができる。

【0154】

次に上述のようにして作成された n 種類の音響モデル $M1$, $M2$, \dots , Mn を用いた音声認識について説明する。

図 13 は実施の形態 2 で用いられる音声認識装置の構成図であり、実施の形態 1 で用いられる音声認識装置（図 7 参照）と異なる点は、雑音パラメータ取得部 13 で取得する雑音パラメータの内容である。

【0155】

この実施の形態 2 では、この雑音パラメータ取得部 13 は、図 10 で説明したように、加工装置動作情報取得部 151、エアコン動作情報取得部 152、ベルトコンベア動作情報取得部 153、検査装置動作情報取得部 154、作業者位置情報取得部 155、窓開度情報取得部 156 などを有している。

【0156】

また、この図 13 の音声認識装置における雑音データ判定部 14 は、これら各情報取得部 151 ~ 156 などからの情報に基づいて、現在の雑音が雑音データ $N1$ ~ $N12$ のうちのどの雑音データに属するかを判定する。

【0157】

たとえば、現在の作業者の位置が $P1$ であって、そのときの加工装置 42 の動作状態が「高速」、エアコン 45 の動作状態が「強風」といった情報を雑音パラメータとして雑音データ判定部 14 が受け取ると、これらの雑音パラメータから

、現在の雑音が雑音データ N1 ~ N12 のどの雑音データに属するかを判定する。この場合、図 11 から現在の雑音が雑音データ N11 に属すると判定される。

【0158】

このように、現在の雑音が雑音データ N11 に属すると判定されたとすると、雑音データ判定部 14 は、その判定結果を雑音除去処理部 16 と音声認識処理部 18 に送る。

【0159】

雑音除去処理部 16 は、雑音データ判定部 14 から、現在の雑音が雑音データ N11 に属するとの情報を受け取ると、入力信号処理部 12 からの雑音重畳音声データに対して最適な雑音除去方法を用いた雑音除去処理を行う。この雑音除去処理は実施の形態 1 で説明したと同様の手法で実現でき、それによって、雑音重畳音声データに対する雑音除去処理がなされる。

【0160】

以上のようにして、マイクロホン 11 から入力されたある時刻の雑音重畳音声データ（作業者の音声とその時点でマイクロホン 11 に入力される雑音からなる）に対する雑音除去処理がなされると、その雑音の除去された雑音除去音声データは音声認識処理部 18 に送られる。

【0161】

音声認識処理部 18 には、雑音データ判定部 14 から現在の雑音がどの雑音データに属するかの情報が与えられており、その雑音データに対応する音響モデルを選択し、その選択された音響モデルと言語モデル 17 を用いて音声認識処理を行う。

【0162】

たとえば、マイクロホン 11 に入力された雑音データは雑音データ N11 に属する雑音であると判定されたとすれば、音声認識処理部 18 では、音響モデルとしては、雑音データ N11 に対応する音響モデル M11 を用いる。

【0163】

この音響モデル M11 は、前述の音響モデル作成方法で説明したように、雑音データ N11 を音声データに重畳させ、その雑音重畳音声データから雑音除去し

て雑音除去音声データを作成し、その雑音除去音声データから作成された音響モデルであるため、作業者の発した音声に重畳された雑音が雑音データN11に属する場合には、その音声にとって最適な音響モデルとなり、認識性能を高めることができる。

【0164】

また、たとえば、現在の作業者の位置がP3であって、そのときの加工装置42の動作状態が「停止」、エアコン45の動作状態が「強風」といった情報を雑音パラメータとして雑音データ判定部14が受け取ると、この雑音データ判定部14は、これらの雑音パラメータから、現在の雑音が雑音データN1～N12のどの雑音データに属するかを判定する。この場合、図1.2から現在の雑音が雑音データN6に属すると判定される。

【0165】

このように、マイクロホン11に入力された雑音データは雑音データN6に属する雑音であると判定されると、音声認識処理部18では、音響モデルとしては、雑音グループN6に対応する音響モデルM6を選択し、その選択された音響モデルと言語モデル17を用いて音声認識を行う。

【0166】

以上のように、この実施の形態2における音声認識装置では、音声コマンドに重畳された雑音が雑音データN1～N12のいずれかに属するかを判定し、それに対応した雑音除去処理手法（音響モデル作成時と同じ雑音除去処理手法）を用いて雑音除去を行い、その雑音除去のなされた音声データ（雑音除去音声データ）に対し、最適な音響モデルを用いて音声認識を行うようにしている。

【0167】

これにより、作業場における作業者の位置やその時々騒音状況に対応した様々な種類の雑音が作業者の音声に重畳されても、その雑音環境下において最適な音響モデルを用いて音声認識することができるので、そのときの作業者の位置や雑音環境下において高い認識性能を得ることができる。

【0168】

なお、本発明は上述の実施の形態に限られるものではなく、本発明の要旨を逸

脱しない範囲で種々変形実施可能となるものである。

たとえば、前述の図 7 および図 13 で示した音声認識装置において、雑音データ判定部 14 は、自動車や作業場の現時点における雑音パラメータを入力することによって現在の雑音が n 種類の雑音データ $N_1 \sim N_n$ のうちのどの雑音データに属するかを判定するようにしているが、この雑音データ判定を行う際、図 14 に示すように、雑音データ判定部 14 には雑音パラメータの他に、音データの重畳された雑音重畳音声データ（デジタル変換後の雑音重畳音声データ）を入力して、この雑音重畳音声データと各種雑音パラメータとによって、現在の雑音が雑音データ $N_1 \sim N_n$ のうちのどの雑音データに属する雑音であるかを判定するようにしてもよい。

なお、この図 14 は実施の形態 1 の図 7 に対応するものであるが、実施の形態 2 の図 13 でも同様ことがいえる。

【0169】

このように、マイクロホン 11 から入力される雑音重畳音声データを雑音データ判定部 14 に入力させることによって、現在の S/N 比をより正確に判別し易くなり、各音響モデル $M_1 \sim M_n$ が S/N 比の大きさも考慮した音響モデルが作成されている場合、現在の S/N 比に応じた最適な音響モデルを選択することができる。より、適正な音声認識を行うことができる。

【0170】

また、雑音パラメータの種類は、前述の各実施の形態で説明した種類に限られるものではなく、その他にも種々用いることができる。なお、音響モデル作成を行うために、実際に自動車を長期間走行させたり、作業場で雑音収集を試みて、統計的手法により分類することによって複数個の雑音データ $N_1 \sim N_n$ を作成したとき、ある雑音パラメータが分類に影響を与えるものでないと判断される場合もあるが、その場合は、音声認識時において、その雑音パラメータは、雑音種類判定部が雑音種類を判定する際の雑音パラメータから除外することができる。

【0171】

また、前述の実施の形態 1 では、乗り物として自動車を例にとって説明したが、自動車に限られるものではなく、たとえば、オートバイなどの二輪車やその他

の乗り物にも適用できることは言うまでもない。

【0172】

同様に、実施の形態2では、工場の作業場を例にとって説明したが、これも工場に限られるものではなく、たとえば、物品の流通センタなど広く適用することができる。

【0173】

また、本発明は以上説明した本発明を実現するための処理手順が記述された処理プログラムを作成し、その処理プログラムをフロッピーディスク、光ディスク、ハードディスクなどの記録媒体に記録させておくこともでき、本発明は、その処理プログラムの記録された記録媒体をも含むものである。また、ネットワークから当該処理プログラムを得るようにしてもよい。

【0174】

【発明の効果】

以上説明したように、本発明の音響モデル作成方法によれば、ある空間内で収集された雑音を分類して複数種類の雑音データを作成し、その複数種類の雑音データを予め用意された標準的な音声データに重畳させて複数種類の雑音重畳音声データを作成し、その複数種類の雑音重畳音声データに対して雑音除去処理を行って、その複数種類の雑音除去音声データから複数種類の音響モデルを作成するようにしているので、その空間内での様々な雑音の種類に対応した最適な音響モデルを作成することができる。

【0175】

また、本発明の音声認識装置は、現在の雑音が複数種類の雑音データのどの雑音データに属する雑音であるかを判別する雑音データ判別を行い、雑音重畳音声に対して、その雑音データ判定結果に基づいた雑音除去を行う。そして、その雑音除去音声に対し、当該雑音データに対応した音響モデルを用いて音声認識を行うようにしている。また、この音声認識装置が用いる複数種類の音響モデルは、上述した音響モデル作成方法によって作成された音響モデルである。これによって、ある空間内に存在する雑音に対し最適な雑音除去処理が行えたとともに、その時点の雑音に対して最適な音響モデルを用いた音声認識が可能となるので、た

例えば、自動車や作業場など特有の雑音環境下で高い認識性能を得ることができる。

【0176】

本また、発明の音声認識装置を有する乗り物は、たとえば、運転者が乗り物自体または乗り物に搭載された機器などの操作や動作設定を行うような場合、その乗り物特有の様々な雑音に適合した音響モデルを用いた音声認識が行えるので、高い認識精度を得ることができ、運転者などが音声で行う操作や動作設定が確実なものとなる。

【図面の簡単な説明】

【図1】 本発明の音響モデル作成方法の大まかな処理手順を説明する図である。

【図2】 本発明の音響モデル作成方法をより詳細に説明する図である。

【図3】 本発明の実施の形態1における雑音データ $N_1 \sim N_n$ の生成処理を説明する図である。

【図4】 ある3種類の雑音パラメータに対応して発生する雑音を長期間かけて収集した雑音データ N を3次元座標上で1つのデータで表した図である。

【図5】 図4の雑音データ N に対し、単純に個々の雑音パラメータごとに分類して得られたそれぞれの雑音グループごとに作成された雑音データを示す図である。

【図6】 図5で示した雑音データに対しある統計的手法によって分類して得られた雑音データを示す図である。

【図7】 本発明の実施の形態1における音声認識装置の構成図である。

【図8】 本発明の音声認識装置を有した乗り物の一例を示す図である。

【図9】 本発明の実施の形態2である工場の作業場の配置を説明する図である。

【図10】 本発明の実施の形態2における雑音データ $N_1 \sim N_n$ の生成処理を説明する図である。

【図11】 本発明の実施の形態2において収集された雑音をある統計的手法によって分類して得られた雑音データを示す図である。

【図 1 2】 図 1 1 を加工装置の 3 つの動作状態のそれぞれに対応する 2 次元断面として表す図である。

【図 1 3】 本発明の実施の形態 2 における音声認識装置の構成図である。

【図 1 4】 図 7 の音声認識装置の変形例を説明する構成図である。

【図 1 5】 従来の音響モデル作成を概略的に説明する図である。

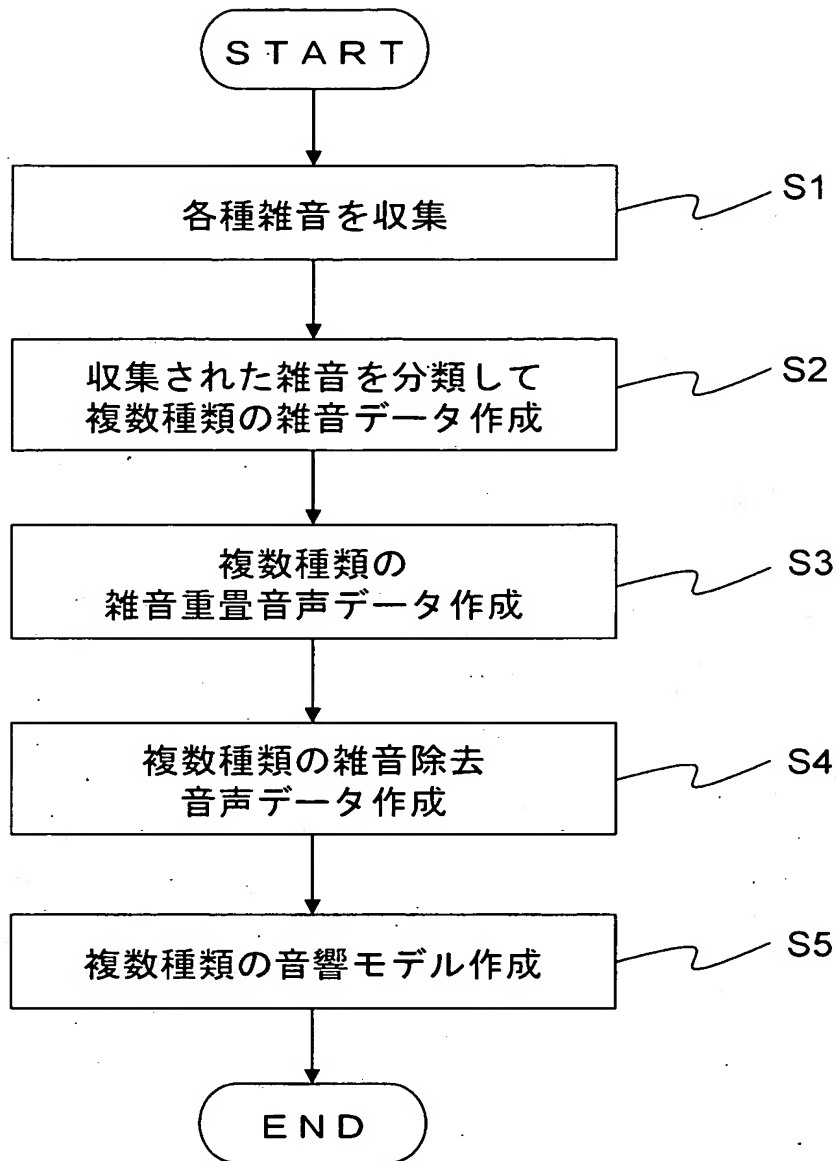
【図 1 6】 図 1 5 で作成された音響モデルを用いた従来の音声認識装置の概略的な構成図である。

【符号の説明】

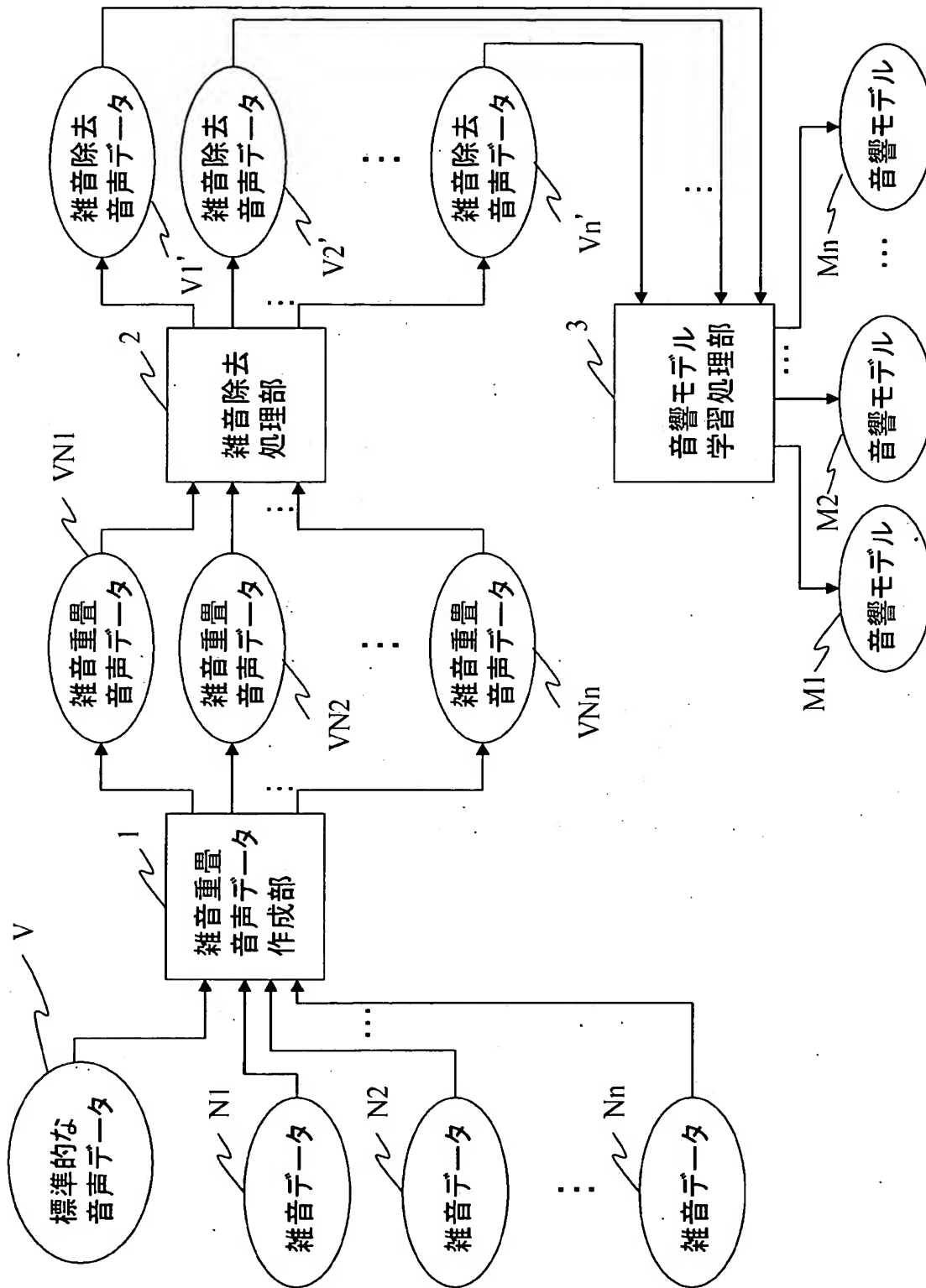
1 雑音重畳音声データ作成部、2 雑音除去処理部、3 音響モデル学習処理部、11 マイクロホン、12 入力信号処理部、13 雑音パラメータ取得部、14 雑音データ判定部、15 雑音除去方法保存部、16 雑音除去処理部、18 音声認識処理部、21 雑音パラメータ記録部、22 雑音記録部、23 雑音分類処理部、N1, N2, ..., Nn 各雑音グループ対応の雑音データ、VN1, VN2, ..., VNn 雑音重畳音声データ、V1', V2', ..., Vn' 雑音除去音声データ、M1, M2, ..., Mn 音響モデル

【書類名】 図面

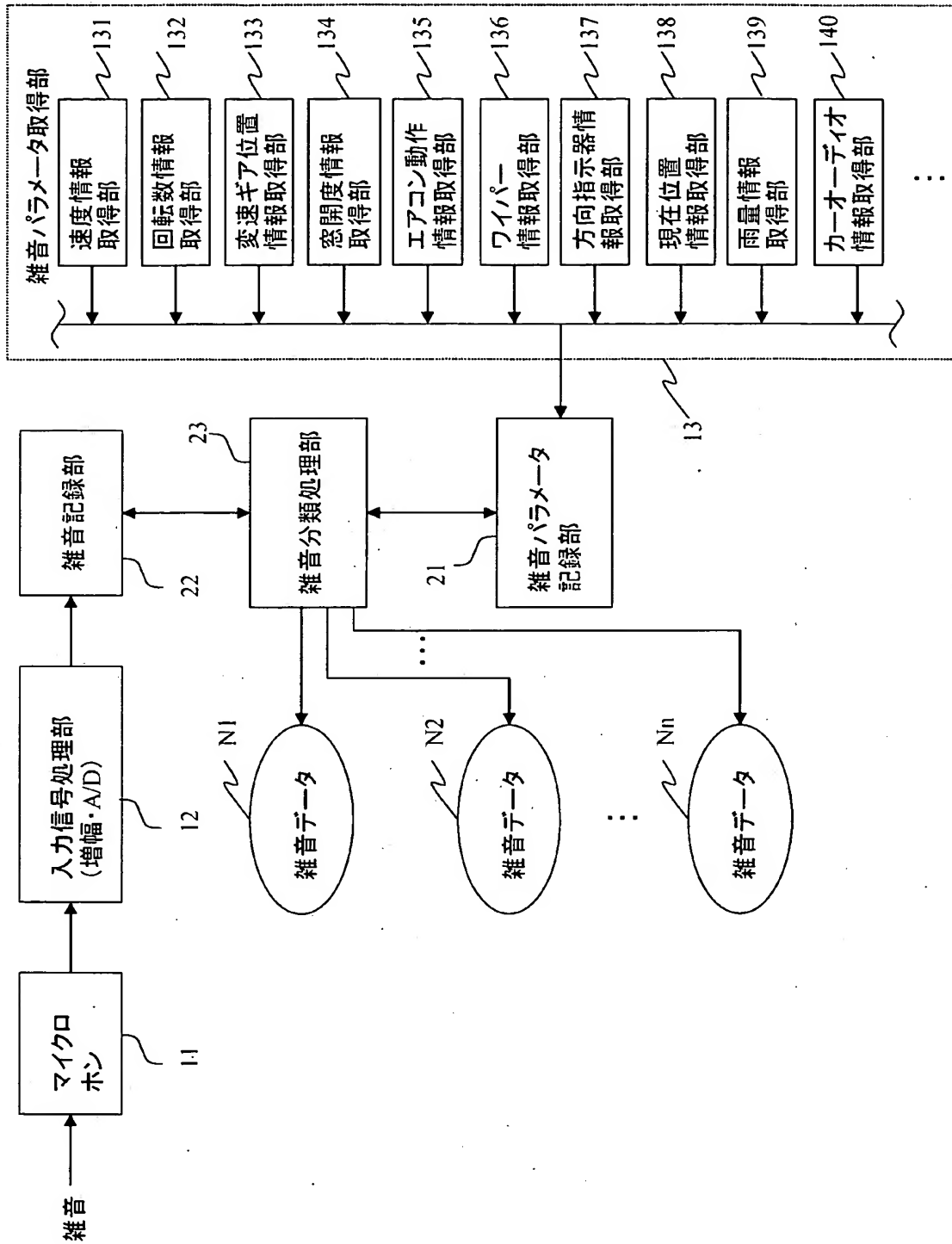
【図 1】



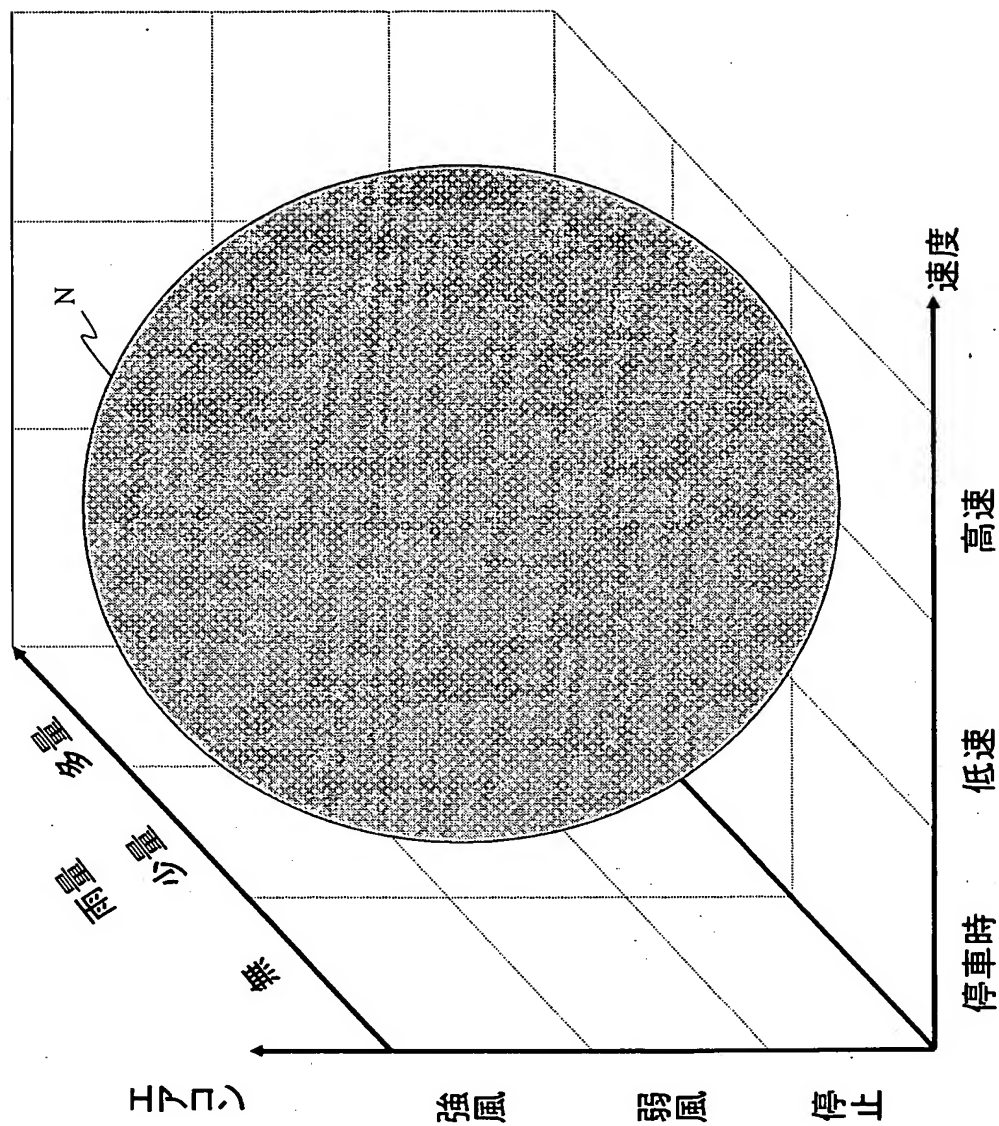
【図2】



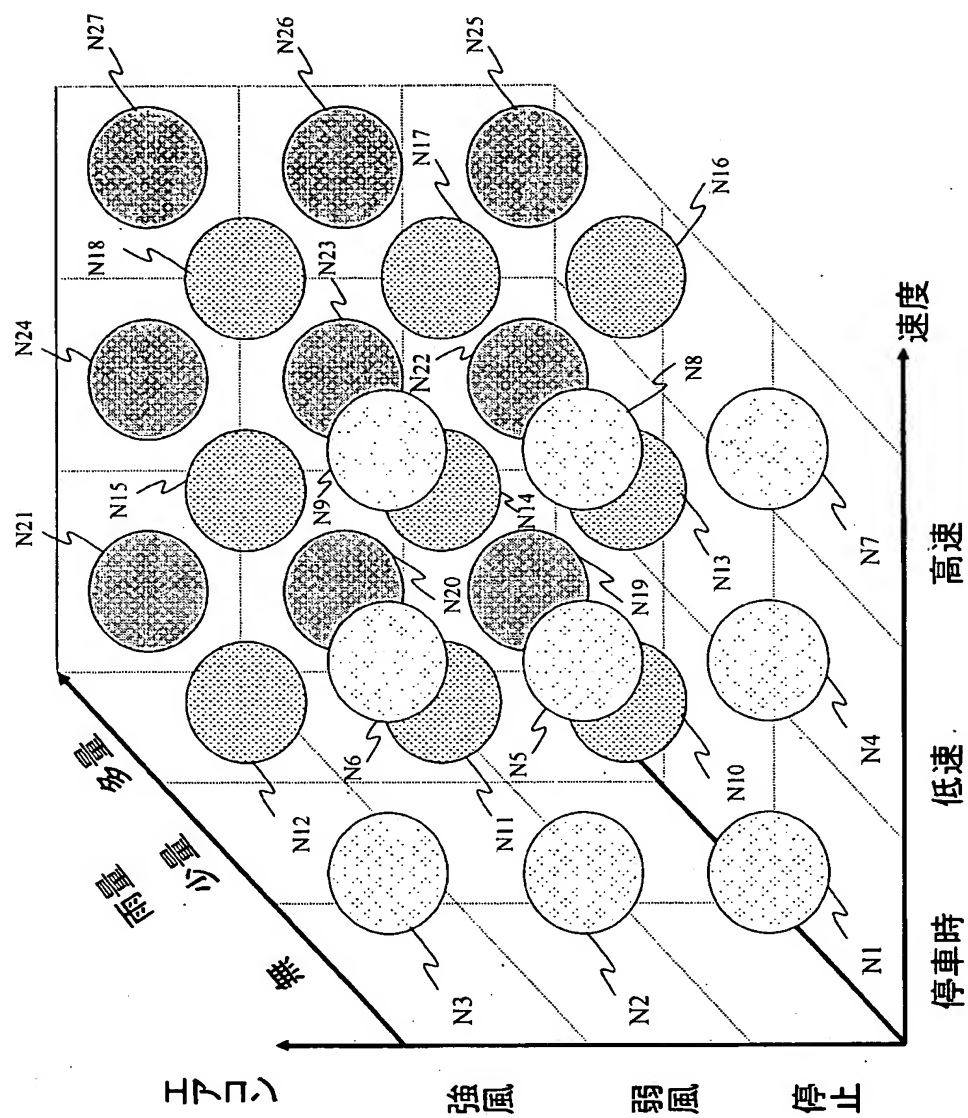
【図 3】



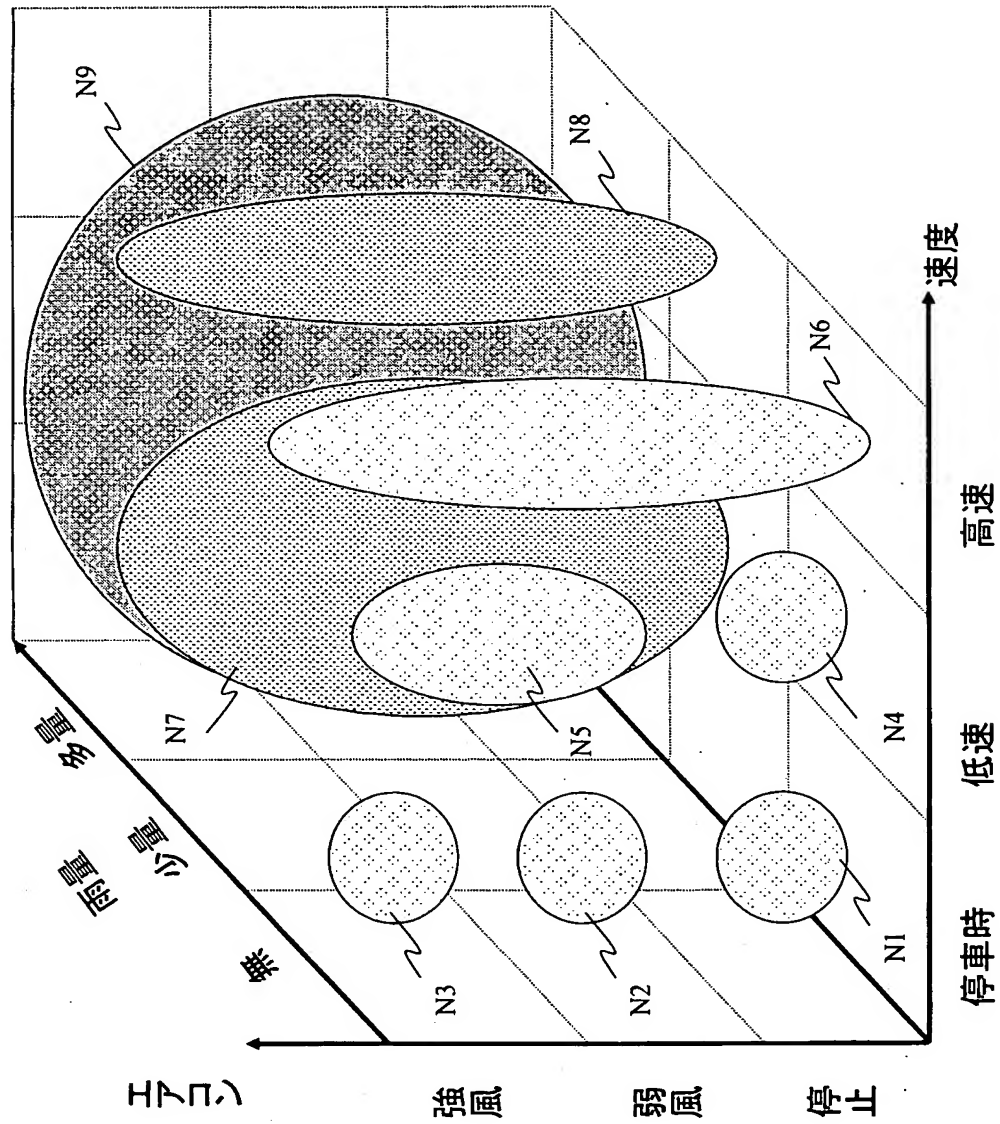
【図 4】



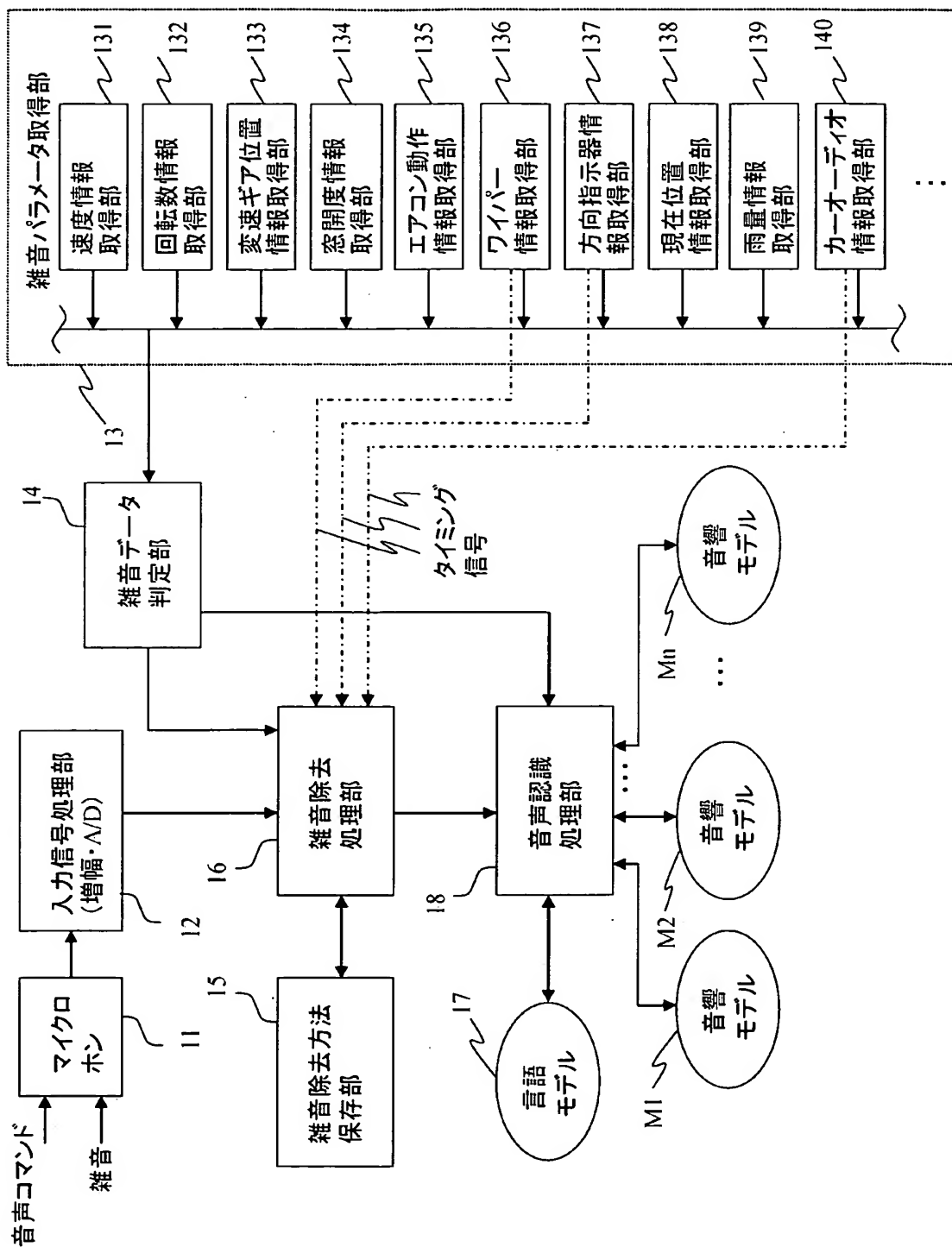
【図 5】



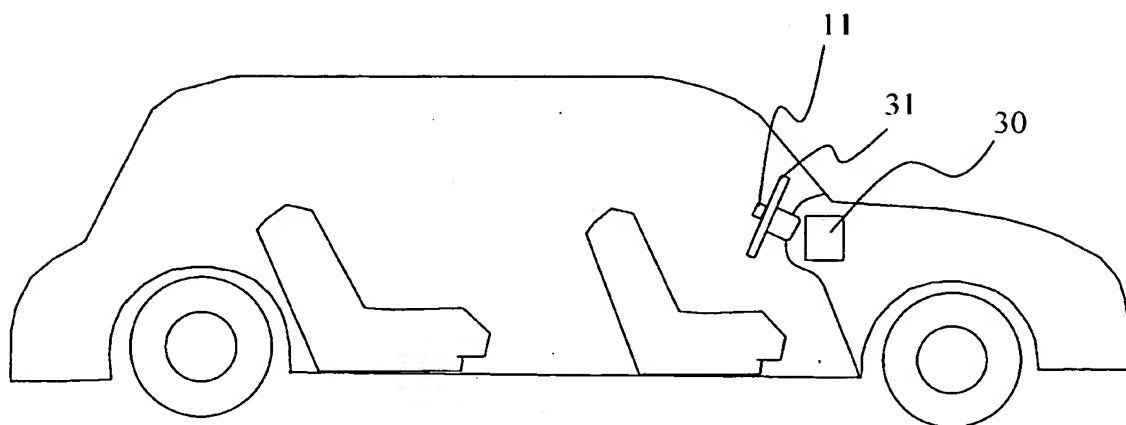
【図 6】



【図 7】

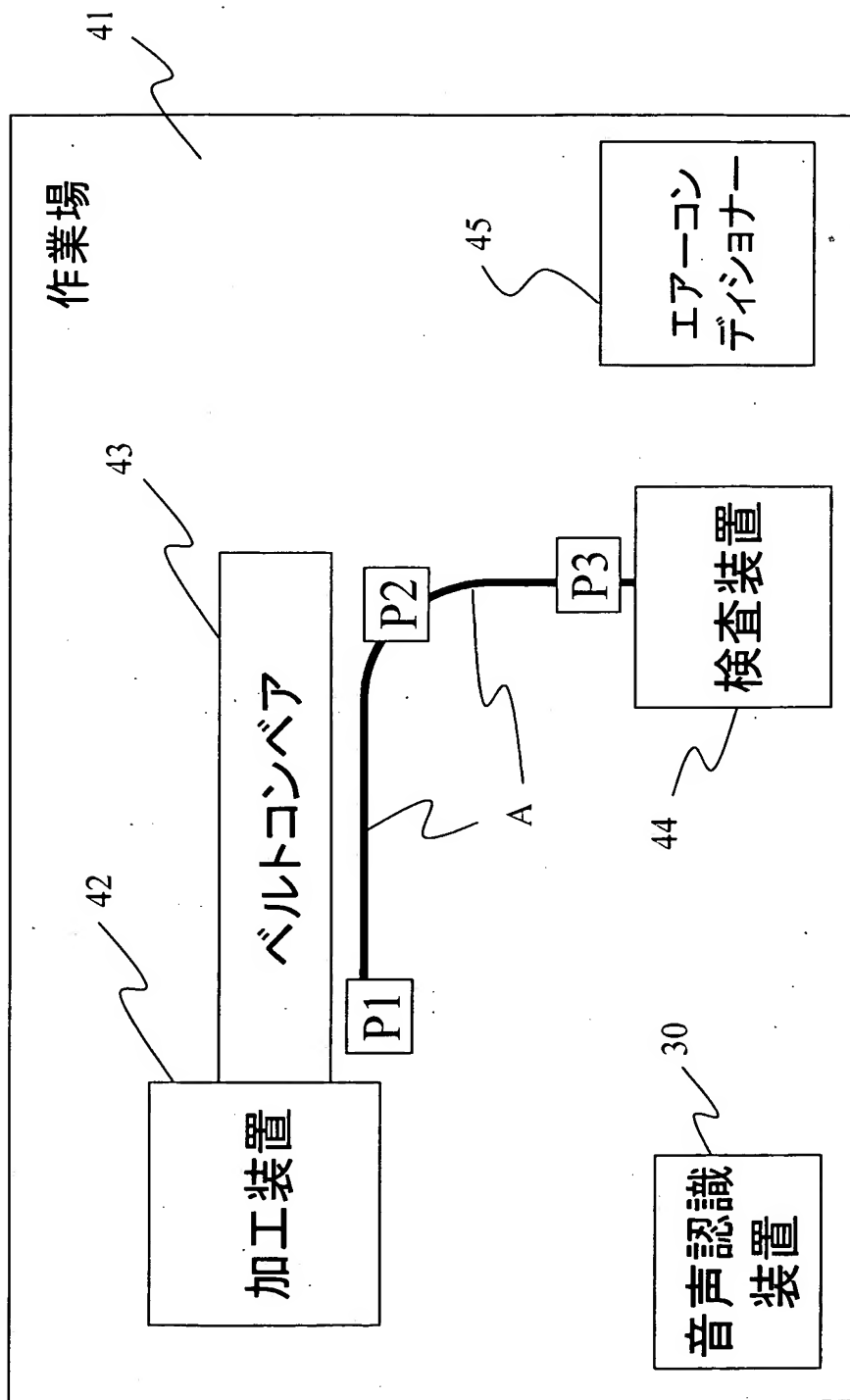


【図 8】

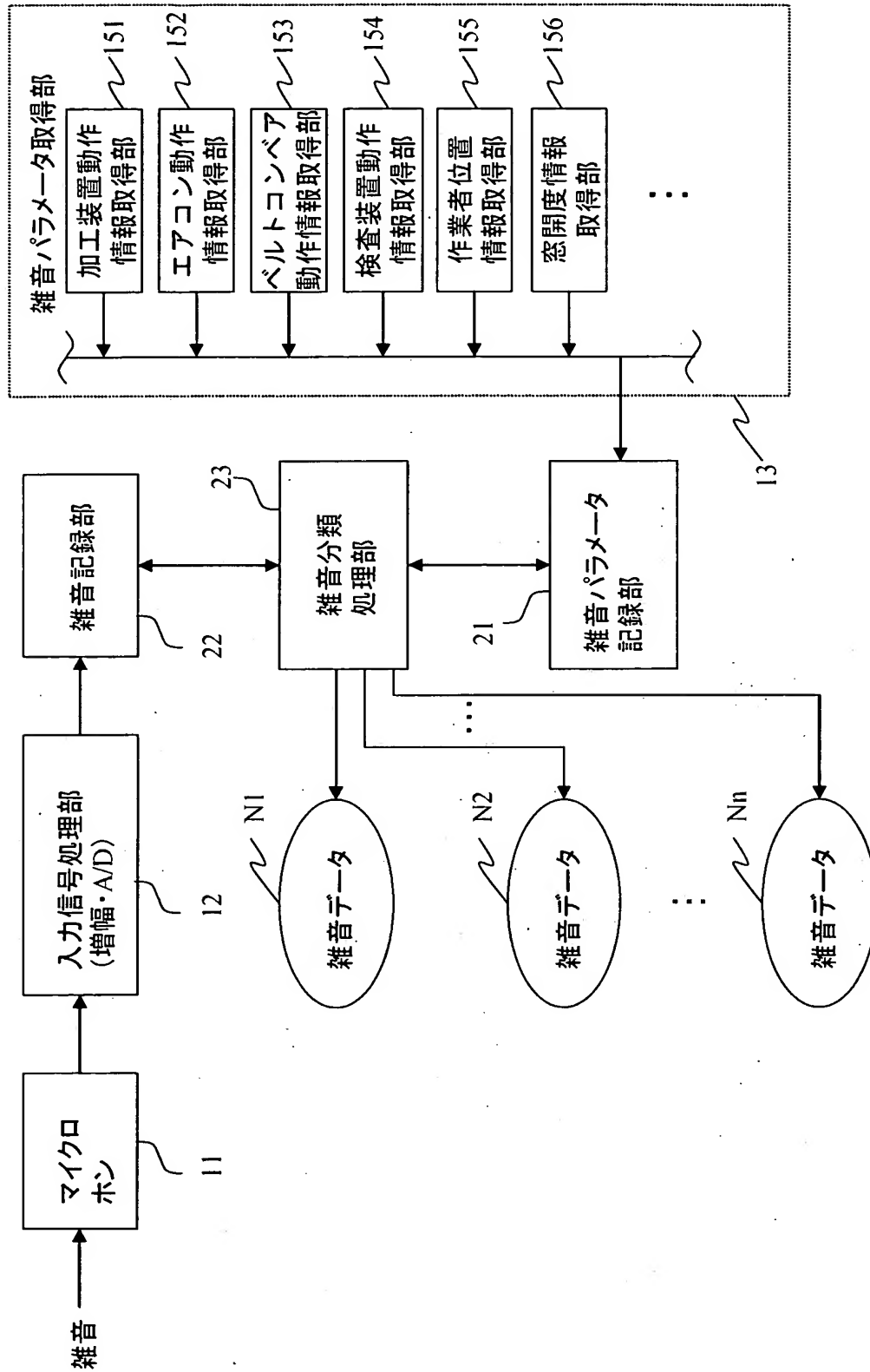


【図 9】

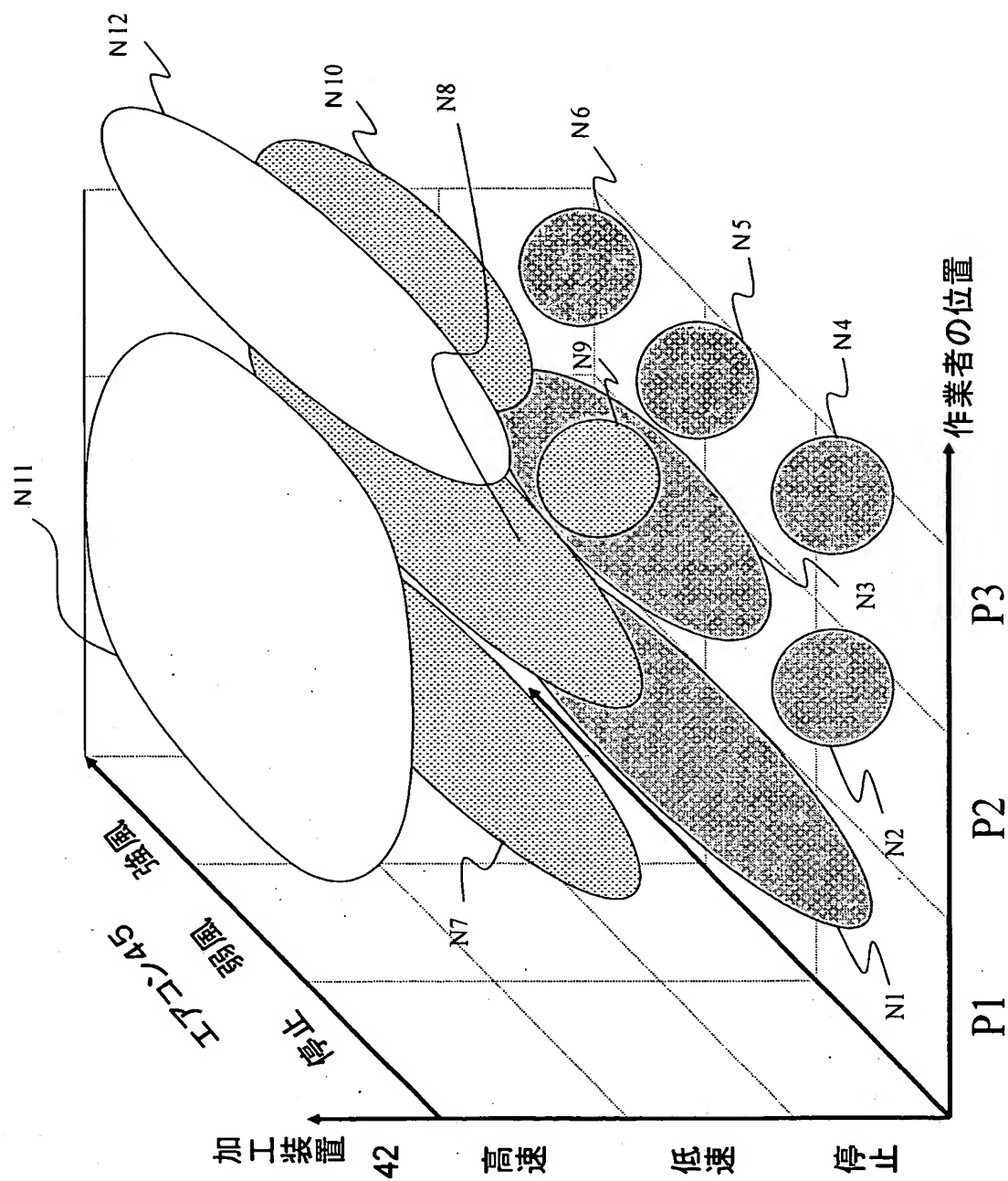
工場内のレイアウト(例)



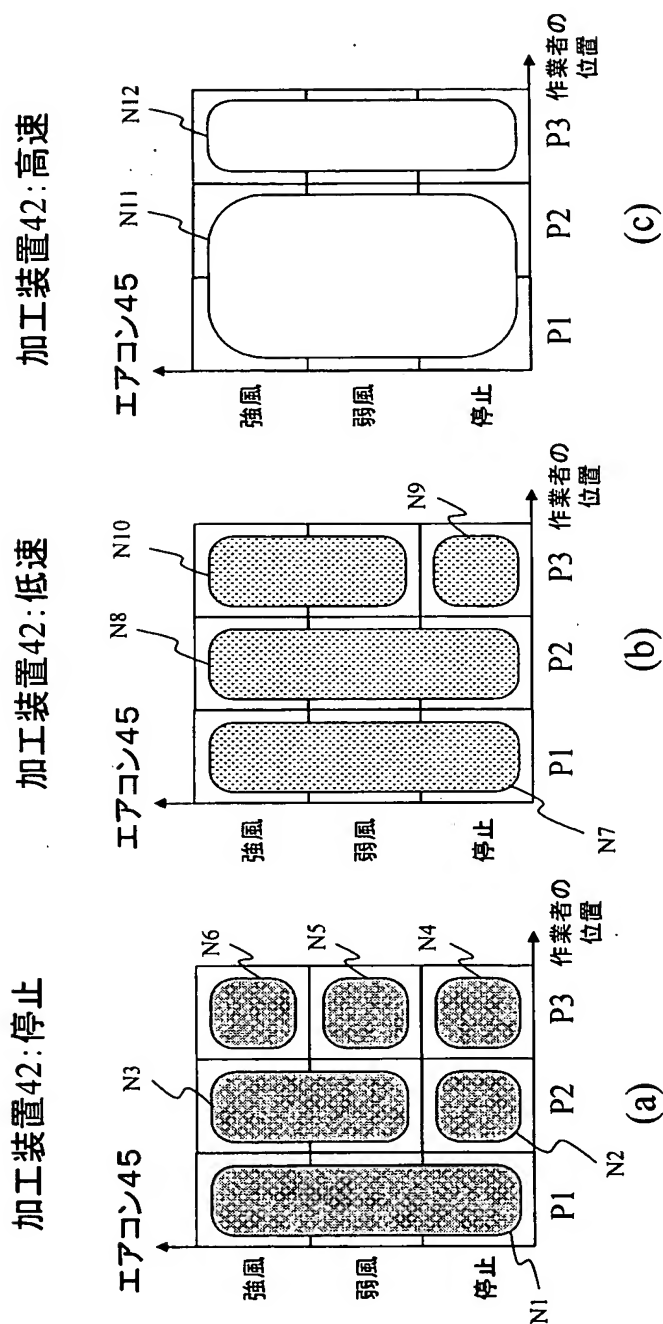
【図 10】



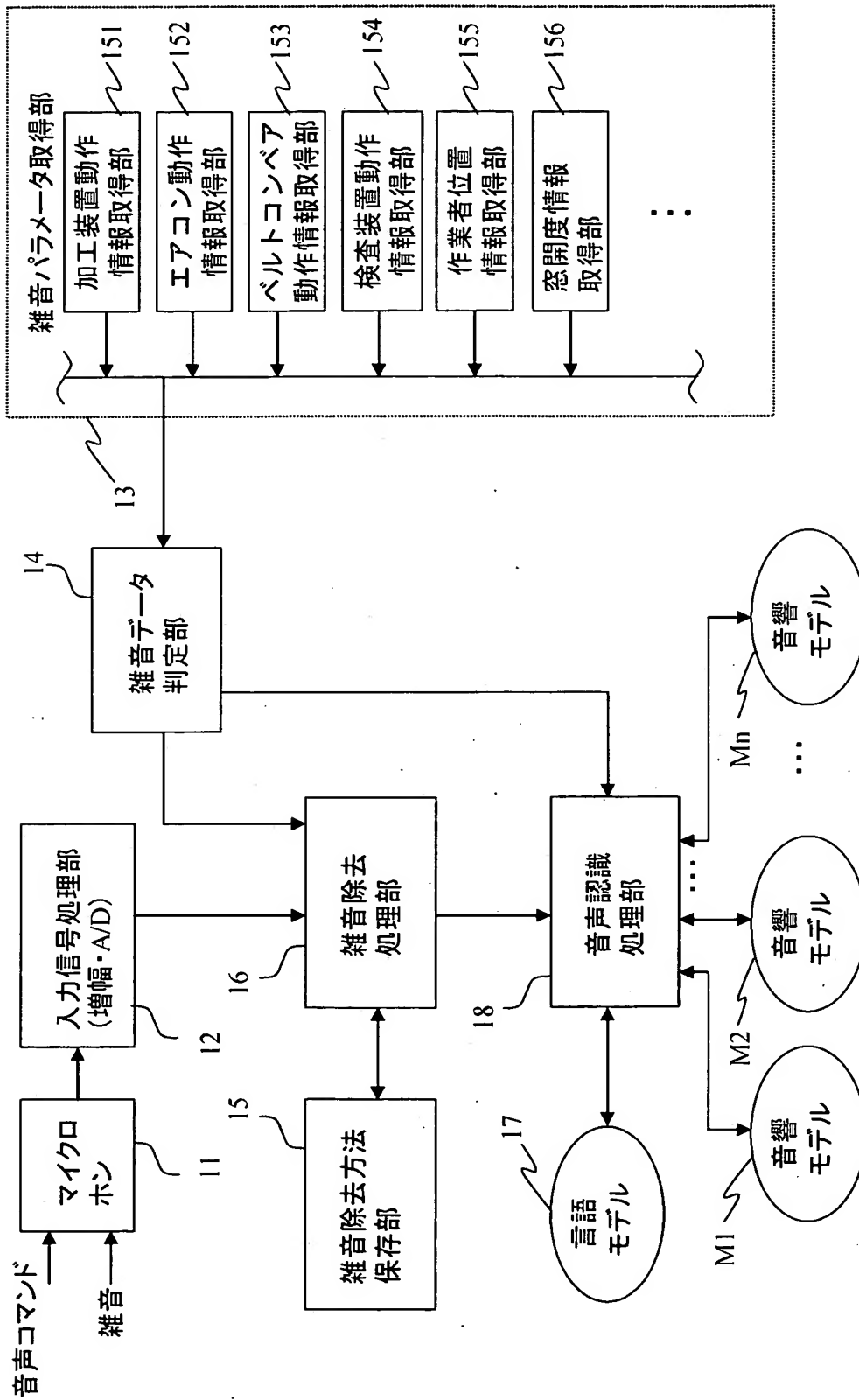
【図 11】



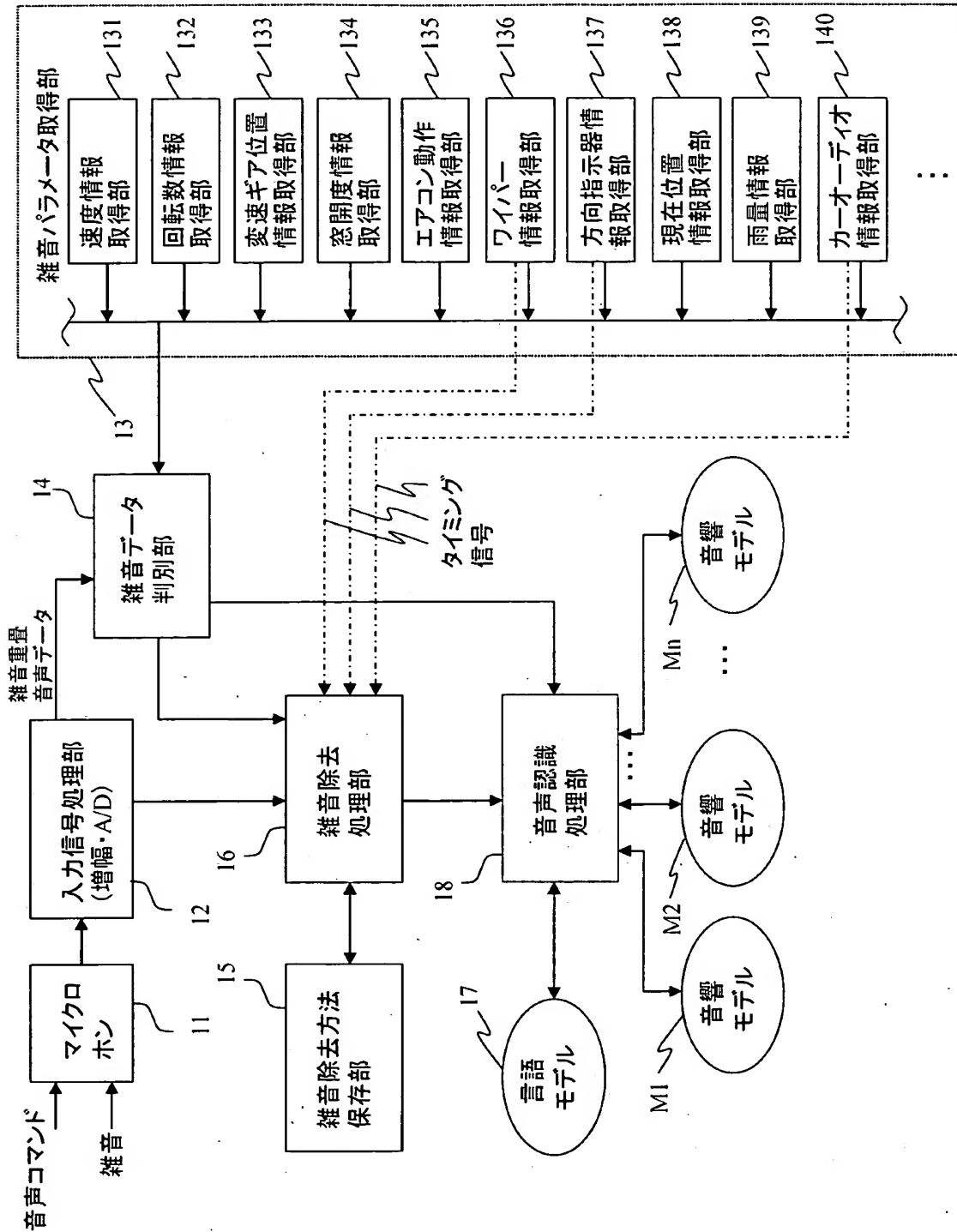
【図 12】



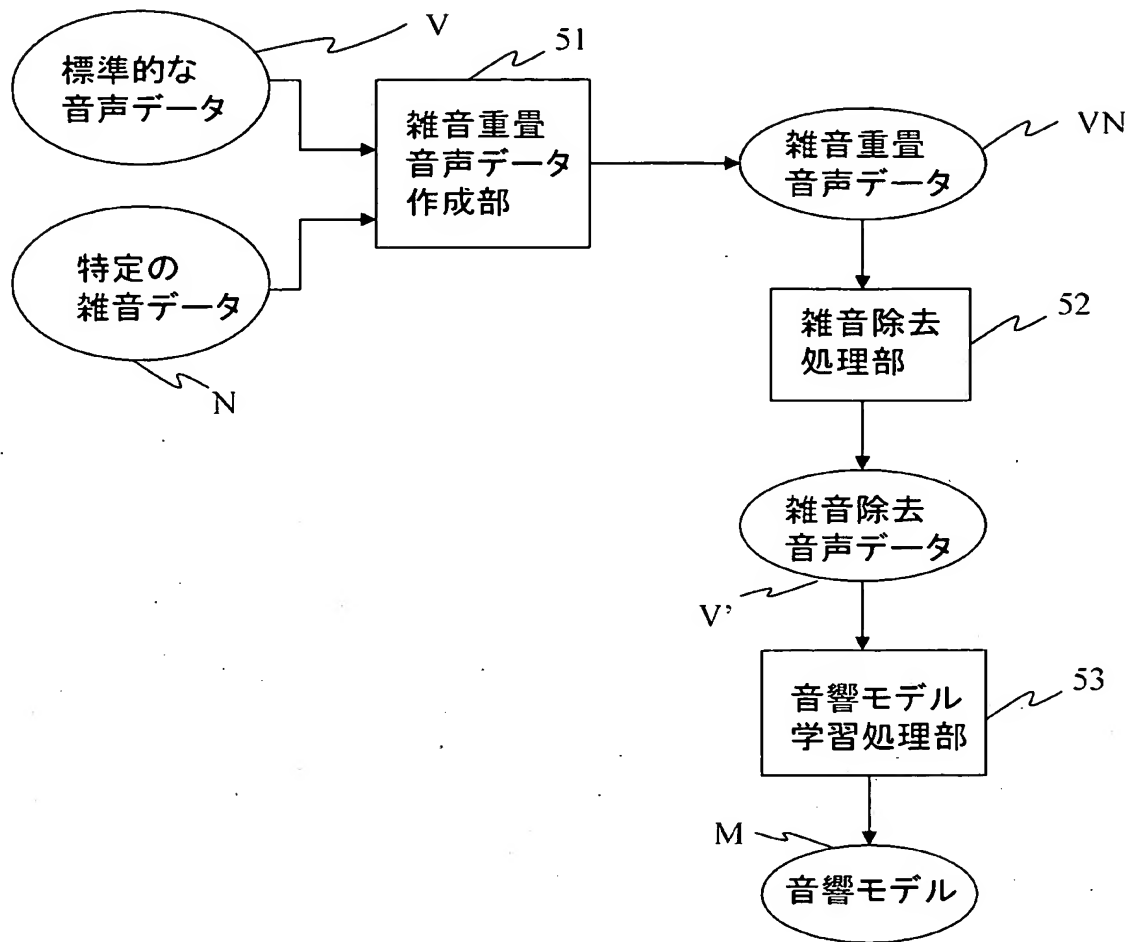
【図 13】



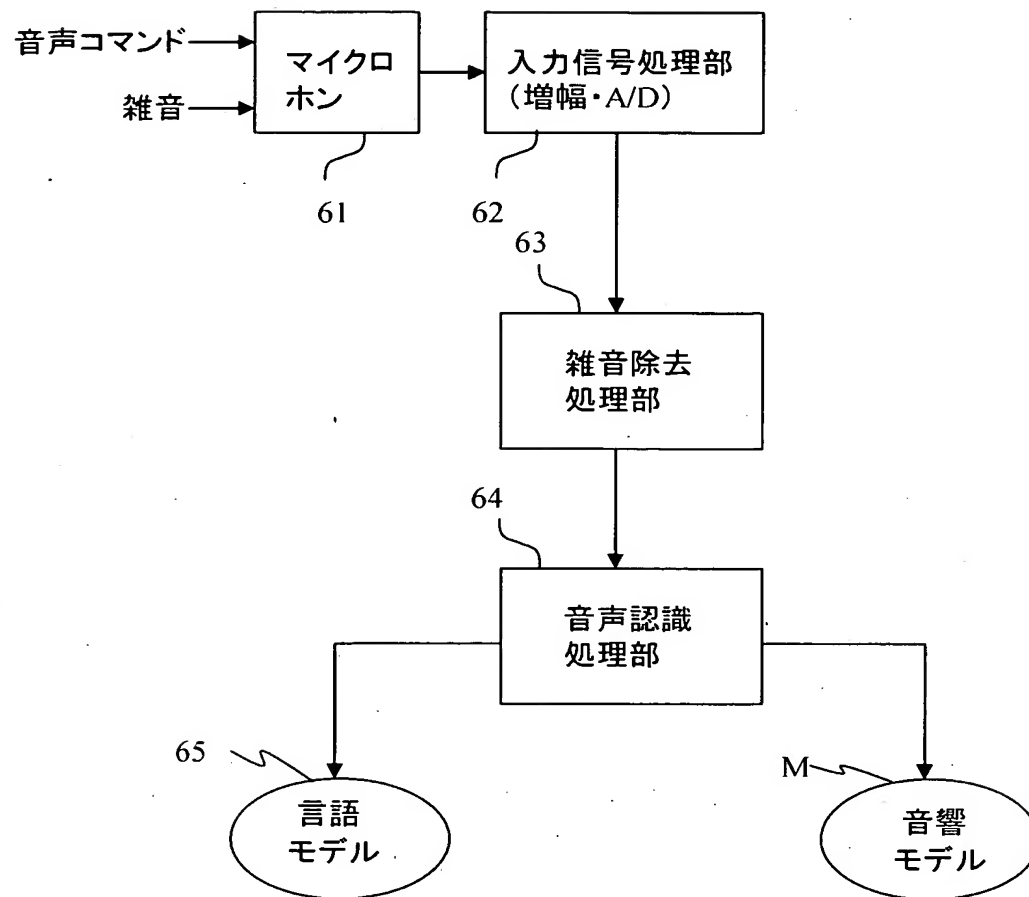
【図 14】



【図 15】



【図 16】



【書類名】 要約書

【要約】

【課題】 自動車内など多種多様な雑音環境下で高い認識性能を得る。

【解決手段】 自動車の走行状況、周囲の環境、搭載機器の動作状態を表すデータを入力し、それによって、現在の雑音が予め分類された n 種類の雑音データのどの雑音データに属するかを判別する雑音データ判定部 14 と、 n 種類の雑音データをそれぞれ標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成したのち、それぞれ雑音除去処理された n 種類の雑音除去音声データに基づいて作成された n 種類の音響モデル $M_1 \sim M_n$ と、マイクロホン 11 から雑音の重畳された雑音重畳音声を入力するとともに、前記雑音種類判定結果を入力し、雑音重畳音声に対して雑音除去を行う雑音除去処理部 16 と、その雑音除去音声に対し、 n 種類の音響モデルのうち雑音データ判定部 14 で判定された雑音種類に対応した音響モデルを用いて音声認識を行う音声認識処理部 18 を有する。

【選択図】 図 7

認定・付加情報

| | |
|---------|------------------|
| 特許出願の番号 | 特願 2003-198707 |
| 受付番号 | 50301189625 |
| 書類名 | 特許願 |
| 担当官 | 第八担当上席 0097 |
| 作成日 | 平成 15 年 7 月 23 日 |

<認定情報・付加情報>

| | |
|-----------|--------------------------------------|
| 【提出日】 | 平成15年 7月17日 |
| 【特許出願人】 | |
| 【識別番号】 | 000002369 |
| 【住所又は居所】 | 東京都新宿区西新宿2丁目4番1号 |
| 【氏名又は名称】 | セイコーエプソン株式会社 |
| 【代理人】 | 申請人 |
| 【識別番号】 | 100095728 |
| 【住所又は居所】 | 長野県諏訪市大和3丁目3番5号 セイコーエプソン株式会社 知的財産本部内 |
| 【氏名又は名称】 | 上柳 雅誉 |
| 【選任した代理人】 | |
| 【識別番号】 | 100107076 |
| 【住所又は居所】 | 長野県諏訪市大和3丁目3番5号 セイコーエプソン株式会社 知的財産本部内 |
| 【氏名又は名称】 | 藤網 英吉 |
| 【選任した代理人】 | |
| 【識別番号】 | 100107261 |
| 【住所又は居所】 | 長野県諏訪市大和3丁目3番5号 セイコーエプソン株式会社 知的財産本部内 |
| 【氏名又は名称】 | 須澤 修 |

特願 2003-198707

出 願 人 履 歴 情 報

識別番号

[000002369]

1. 変更年月日

1990年 8月20日

[変更理由]

新規登録

住 所

東京都新宿区西新宿2丁目4番1号

氏 名

セイコーエプソン株式会社